



PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 5 :</b> C12N 15/11, 15/79, 15/80 C07K 13/00		<b>A1</b>	<b>(11) International Publication Number:</b> WO 91/06644 <b>(43) International Publication Date:</b> 16 May 1991 (16.05.91)
<b>(21) International Application Number:</b> PCT/AU90/00530		<b>(74) Agent:</b> F.B. RICE & COMPANY; 28A Montague Street, Balmain, NSW 2041 (AU).	
<b>(22) International Filing Date:</b> 2 November 1990 (02.11.90)			
<b>(30) Priority data:</b> PJ 7187 2 November 1989 (02.11.89) AU		<b>(81) Designated States:</b> AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CM (OAPI patent), DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL, NL (European patent), NO, RO, SD, SE, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US.	
<b>(71) Applicant (for all designated States except US):</b> MACQUARIE UNIVERSITY [AU/AU]; Balaclava Road, North Ryde, NSW 2113 (AU).		<b>Published</b> <i>With international search report.</i>	
<b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only) :</b> SLADE, Martin, Basil [AU/AU]; 40 Conrad Street, East Ryde, NSW 2113 (AU). CHANG, Andy, Chin, Meng [NZ/AU]; 7/19 Bowden Street, Harris Park, NSW 2150 (AU). WILLIAMS, Keith, Leslie [AU/AU]; 25 Kara Street, Lane Cove, NSW 2066 (AU).			

**(54) Title:** IMPROVED PLASMID VECTORS FOR CELLULAR SLIME MOULDS OF THE GENUS DICTYOSTELIUM

**(57) Abstract**

The present invention relates generally to the fields of molecular biology and the production of recombinant protein using cellular slime moulds of the genus *Dictyostelium*. Most particularly, the present invention relates to novel strains of the genus *Dictyostelium*, recombinant plasmids for use with strains of the genus *Dictyostelium*, and polypeptides which facilitate the extrachromosomal replication of such plasmids in strains of the genus *Dictyostelium*. In particular, the present invention provides a polypeptide which facilitates the extrachromosomal replication of a recombinant plasmid in *Dictyostelium* spp in which the recombinant plasmid includes an origin of replication derived from a *Ddp2*-like plasmid but which lacks functional genes for extrachromosomal replication in wild type *Dictyostelium* spp. The extrachromosomal replicating plasmid constructed in accordance with the present invention are suitable for carrying a wide variety of genes and promoter sequences for control production of recombinant proteins by the biotechnology industry.

***FOR THE PURPOSES OF INFORMATION ONLY***

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MR	Mauritania
BE	Belgium	GA	Gabon	MW	Malawi
BF	Burkina Faso	GB	United Kingdom	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CI	Côte d'Ivoire	LK	Sri Lanka	TD	Chad
CM	Cameroon	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America

IMPROVED PLASMID VECTORS FOR CELLULAR SLIME MOULDS OF THE GENUS DICTYOSTELIUM

Field of the Invention

The present invention relates generally to the fields 5 of molecular biology and the production of recombinant proteins by the biotechnology industry. More particularly, the present invention relates to novel strains of the genus Dictyostelium, recombinant plasmid vectors for use with strains of the genus Dictyostelium, and polypeptides 10 which facilitate the extrachromosomal replication of such plasmids in strains of the genus Dictyostelium. Such extrachromosomally replicating plasmids, constructed with the art disclosed in this invention, are suitable for carrying a wide variety of genes and promoter sequences 15 for the controlled production of recombinant proteins by the biotechnology industry.

BACKGROUND ART

As is well known in the art, genetic information is encoded on double stranded DNA molecules according to the 20 sequence of four nucleotides containing different bases, adenine (A), thymine (T), cytosine (C) and guanine (G). Blocks of DNA sequences flanking genes often control gene activity by binding regulatory proteins and acting as 25 recognition signals for enzymes of the cells biosynthetic machinery. Thus each cell contains a web of regulatory molecules which, by binding to specific DNA sequences, control gene activity. Other DNA sequences have crucial functions related to the control of DNA synthesis and partitioning of DNA into separate cells during cell 30 division. These functions must be present on every DNA molecule in every cell or the DNA will be lost within a few cell generations.

Plasmids are usually circular DNA molecules 35 possessing DNA sequences allowing them to replicate independently from chromosomal DNA. The DNA sequence

block where the replication of plasmid DNA is initiated is commonly called the "origin of replication" and the ability to replicate independently from chromosomal DNA is referred to as "extrachromosomal" replication.

5        Molecular biologists have developed techniques for cutting DNA molecules into fragments using sequence specific restriction enzymes, purifying the fragments and rejoining them in a different order. If one of the fragments of DNA used contains an origin of replication  
10      from an E. coli plasmid, the DNA can be inserted (transformed) into E. coli where it will replicate as a plasmid and can be produced in relatively large quantities. These techniques mean that genes from one organism, for example a human gene, can be flanked by  
15      regulatory DNA sequences from another organism, for example the bacterium E. coli, causing the human gene to be active in E. coli under entirely different regulatory controls. If the plasmid in question is constructed to include a second origin of replication allowing  
20      replication in a separate host cell, for example a mouse cell line, the gene can easily be transferred to the second host cell. Such a plasmid containing origins of replication for more than one host is commonly called a "shuttle vector". Plasmids are usually constructed to  
25      contain selectable markers, which are usually genes that confer antibiotic resistance or a metabolic advantage on the host cell to allow cells containing the plasmid to be distinguished from cells that have not received any plasmid during the transformation. Selectable marker  
30      genes must be flanked by appropriate DNA sequences to permit gene activity in the required host cell. It is possible to insert a plasmid into a host cell where it will be unable to replicate and so the only cells that survive the selection procedure will be those with the  
35      plasmid inserted into the host's chromosomal DNA. Such a

plasmid without an appropriate origin of replication is called an "integrating plasmid".

A cell produces polypeptides and proteins by initially making a messenger RNA copy of the gene, a 5 process called transcription which is under the control of the flanking DNA sequences as summarised above. The cellular biosynthetic machinery then reads (translates) the RNA sequence in three nucleotide groups called codons which specify the amino acids to be 10 incorporated into the polypeptide chain. The genetic code and mechanism of protein synthesis is very similar in all organisms so molecular biology techniques can be used to construct plasmid vectors to produce recombinant proteins in many different host cells irrespective of the source of 15 the original gene. However, different host cells may process the protein in different ways so it may, for example, be folded incorrectly or cleaved by protease enzymes. Most importantly, eukaryotic cells differ from bacteria by frequently linking further chemical structures 20 onto their proteins, a process called "post-translational modification". The chemical structures linked to eukaryotic proteins may include several types of oligosaccharide chains, glycolipids, lipids, sulphate and phosphate groups, all of which may affect the physical and 25 biological properties of the molecule. Common effects of these post-translational modifications include increased resistance to proteolysis, altered immunogenicity, altered in vivo clearance and uptake by different cell types.

Post-translational modifications frequently occur on 30 proteins that are secreted from cells or are present on cell membranes. Such proteins include a wide variety of soluble proteins that mediate inter-cellular interactions, blood proteins and cell surface receptors and so are of considerable interest to the pharmaceutical industry as 35 either the targets for drug research or for in vivo

- 4 -

administration as therapeutic drugs in their own right. Since post-translational modifications may substantially alter the biological activity of such proteins (for example, tissue plasminogen activator (Ezzell, 1988, 5 Nature 333, 383)), it is a goal of the biotechnology industry to produce each protein with a range of different modifications, both those that occur naturally and new modifications such as truncated oligosaccharide chains. However, proteins with post-translational modifications 10 can only be produced in eukaryotic hosts and only a few eukaryotes have been used industrially. Mammalian tissue culture, for example Chinese Hamster Ovary Cells, is usually able to produce proteins with post-translational modifications similar to the natural protein, but is very 15 expensive since these cells frequently require serum components in their growth media, have a slow growth rate and are relatively difficult to grow in large fermentors. Consequently, simple eukaryotes such as insect cells infected with baculovirus or yeast cells have been used to 20 produce proteins with some post-translational modifications at a considerably lower cost. However, no one host is suitable for all recombinant proteins or can produce more than a few of the wide range of desirable post-translational modifications.

25 Dictyostelium has some advantages as a host for the production of low cost recombinant proteins with post-translational modifications (reviewed by Glenn & Williams, 1988, Australian J. Biotech. 1(4), 46-56). These include the production of N-linked glycosylation 30 indistinguishable from the mammalian "high mannose form" and a wide variety of other structures including phosphatidyl-inositol-glycan tails. It is possible to alter the post-translational modifications produced by Dictyostelium by either using a range of mutant cultures 35 which produce altered glycan structures or by simply

- 5 -

harvesting the Dictyostelium cells at different stages of the life cycle. A considerable body of scientific literature is available on the culture and genetics of Dictyostelium (Spudich J. Ed. (1987) Methods in Cell Biology Vol. 28, Academic Press, London). Dictyostelium has a number of characteristics suitable for use in the production of recombinant proteins in fermenters since they grow rapidly (4-10 hour cell cycle) and reach high densities (around 50 million cells per ml) in a nutrient medium. For some purposes, the ability of Dictyostelium to grow on a lawn of bacteria on a simple nutrient medium provides a remarkably simple and cheap culture technique when compared with mammalian or even insect tissue culture.

Dictyostelium strains are known to possess at least thirteen different plasmids (Farrar & Williams (1988) Trends in Genetics 4, 343-348), but only Ddp1, Ddp2 and pDG1 have been studied in detail. Plasmid pDG1 is very unstable when cloned in E. coli (Orii et al (1989) Nucleic Acids Research 17, 1395-1408) so most constructions of shuttle vectors have used sequences from either Ddp1 or Ddp2. Plasmid Ddp1 is 12.3 Kb in size, but Ahern et al (Nucleic Acids Research (1988) 16, 6825-6837) showed that a vector containing a selectable marker (G418) resistance and only 2.2 Kb of Ddp1 was able to replicate extrachromosomally in D. discoideum. However, but the copy number per cell of this truncated plasmids lowered from the 150 characteristic of the parent plasmid to only 10-15 copies per cell. It is probable that this low copy number plasmid may not segregate efficiently at cell division and so may be unstable in the absence of continuous selection with the antibiotic G418. Incorporation of additional Dictyostelium DNA into such plasmids based on the Ddp1 origin of replication prevents them being maintained extrachromosomally (Gurniak et al,

- 6 -

(1990) Current Genetics 17, 321-325.) so they are unsuitable for use in the biotechnology industry.

The practical application of plasmids constructed from sections of Ddp2 has been limited by technical difficulties. The majority of techniques used in molecular biology are designed for use in the bacterium E. coli so the manipulation of Dictyostelium DNA requires it to be cloned into a vector capable of replication in E. coli. Consequently, research on Ddp2 has concentrated on the construction of recombinant "shuttle vectors" containing sequences allowing replication in both E. coli and Dictyostelium spp. Plasmid pMUW111 illustrates a shuttle vector that the present inventors have constructed (Fig. 4), which contains a 4.139 Kb Hind III - ScaI restriction fragment of Ddp2. This is close to the minimum amount of Ddp2 which can maintain extrachromosomal replication in wild type strains of Dictyostelium. Leiting and Noegel (1988 Plasmid 20, 241-248) have used a similar 4.0 Kb fragment of Ddp2 with approximately 300bp deleted close to the Xho I restriction site to construct a 9.6 Kb shuttle vector called pnDE1. However, despite containing minimal sections for the extrachromosomal replication of Ddp2, both these shuttle vectors (pMUW111 and pn DE1) suffer from problems of instability when maintained in E. coli. This is consistent with the Ddp2 DNA containing sequences that are unstable in E. coli. This problem can be mitigated by the use of host strains which lack exo-nuclease I and have low plasmid copy number (eg strain CES 201), but such hosts frequently present problems in preparing sufficient plasmid DNA for gene cloning experiments and for transforming back into Dictyostelium.

The necessity of using pieces of Ddp2 DNA approximately 4 Kb long to construct shuttle vectors also raises problems with regard to the final size of the

plasmid. The shuttle vector must contain selectable markers for both hosts together with appropriate promoter and termination sequences. These sequences comprise nearly 50% of the size of plasmids pMUW111 and pnDe1. In 5 addition, to be of any practical use a shuttle vector must be capable of carrying additional DNA containing a gene to be expressed in Dictyostelium together with appropriate controlling sequences. These additional sequences are likely to amount to a minimum of at least 2 Kb of DNA, 10 bringing the total plasmid size to around 12 kilobase pairs. Increasing the size of the plasmid to over 10 Kb decreases its stability, a factor of considerable importance for the commercial production of recombinant proteins where, in order to avoid contamination of the 15 product, regulatory authorities do not permit the use the antibiotic selection to ensure plasmid maintenance while cells are grown for extended periods. A large plasmid also raises difficulties since fewer restriction enzymes will cut the plasmid at only one position, the most 20 suitable sites for genetic manipulations.

Shuttle vectors capable of being easily manipulated in E. coli and transferred back into Dictyostelium spp. are an essential pre-requisite for realising the potential of Dictyostelium in biotechnology. The present inventors 25 have discovered means by which such vectors containing sections of Ddp2 smaller than 4 Kb can be constructed.

The present inventors have elucidated the full nucleotide sequence of the plasmid Ddp2 and have determined that a portion of this sequence encodes a gene 30 designated Rep. The present inventors have shown that the presence of a polypeptide encoded by the Rep gene is essential for extrachromosomal replication of the Ddp2 plasmid.

#### Disclosure of the Invention

35 Accordingly, in a first aspect the present invention

consists in a polypeptide which facilitates the extrachromosomal replication of a recombinant plasmid in Dictyostelium spp, the recombinant plasmid including an origin of replication derived from a Ddp2-like plasmid, 5 but lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

In a preferred embodiment of this aspect of the present invention the recombinant plasmid includes an origin of replication derived from plasmid Ddp2.

10 In a preferred embodiment of this aspect of the present invention the polypeptide has an amino acid sequence substantially as shown in Figure 2.

15 In a further preferred embodiment of this aspect of the present invention the polypeptide is encoded by a DNA sequence substantially as shown in Figure 1 from nucleotide 2378 to nucleotide 5038.

20 As used herein the phrase "Ddp2-like plasmid" is intended to cover plasmids having similar structure and similar functional regions to plasmid Ddp2. One example of such a Ddp2-like plasmid is plasmid pDG1.

25 In a second aspect the present invention consists in a recombinant plasmid vector, said vector being characterised in that it includes an origin of replication derived from plasmid Ddp2 or plasmid pDG1 and that it lacks functional genes for extrachromosomal replication in wild type Dictyostelium.

30 In a third aspect the present invention consists in a recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1 to nucleotide 2436 or a subsection thereof, and lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

35 In a fourth aspect the present invention consists in a recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1153 to

- 9 -

- nucleotide 1775 or a subsection thereof, and lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

In a fifth aspect the present invention consists in a 5 recombinant plasmid vector containing the DNA sequence TGTCA TGACA but lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

In a sixth aspect the present invention consists in a 10 recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1 to nucleotide 3241 or a portion thereof and lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

15 It is presently preferred that the recombinant plasmid vector includes a heterologous DNA sequence(s) encoding a desired polypeptide, a promoter sequence(s) that controls the expression of the heterologous DNA sequence(s), and preferably a sequence(s) including a 20 selectable marker.

In a preferred embodiment of the present invention the recombinant plasmid vector includes a DNA sequence encoding a polypeptide and regulatory sequences for secretion of the desired polypeptide.

25 In a further preferred embodiment of the present invention the recombinant plasmid vector includes an expression cassette comprising a promoter DNA sequence derived from the Dictyostelium Actin 15 gene, a DNA sequence encoding the secretion signal peptide sequence of 30 the D19 gene which encodes the protein PSA and a DNA sequence for RNA polyadenylation signal derived from the Actin 15 gene.

In a further preferred embodiment of the present invention, the recombinant vector includes the sequence of 35 plasmid pMUW102, plasmid pMUW130 or plasmid pMUW1530 and a

- 10 -

heterologous DNA sequence encoding a desired polypeptide together with DNA sequences enabling the expression of the sequence encoding the desired polypeptide.

In a seventh aspect, the present invention consists  
5 in a recombinant strain of Dictyostelium, the recombinant strain being characterised in that the strain includes a gene encoding a polypeptide which facilitates the extrachromosomal replication of a recombinant plasmid, the recombinant plasmid including an origin of replication  
10 derived from plasmid Ddp2 but lacking the functional gene for extrachromosomal replication in wild type  
Dictyostelium.

In a preferred embodiment of the present invention the recombinant plasmid includes an origin of replication  
15 derived from plasmid Ddp2, and is more preferably the recombinant plasmid of one of the second to sixth aspects of the present invention.

The gene encoding the polypeptide which facilitates the extrachromosomal replication of the recombinant  
20 plasmid may be present in a chromosome of the recombinant strain of Dictyostelium or carried on a second plasmid, the second plasmid lacking an origin of replication derived from Ddp2. It is, however, presently preferred that the gene encoding the polypeptide is carried on a  
25 chromosome.

It is presently preferred that the recombinant strain of Dictyostelium has included within a chromosome the Rep gene.

In a further preferred embodiment of the present  
30 invention the chromosome of the recombinant strain of Dictyostelium includes a sequence substantially as shown in Figure 1 from nucleotide 1885 to nucleotide 5292.

In a further preferred embodiment of the present invention the recombinant strain of Dictyostelium harbors  
35 a recombinant plasmid, the recombinant plasmid including

- 11 -

an origin of replication derived from plasmid Ddp2 or plasmid pDG1, and preferably a DNA sequence encoding a desired polypeptide together with a DNA sequence enabling the expression of the sequence encoding the desired 5 polypeptide, but lacking functional genes for extrachromosomal replication in wild type Dictyostelium.

In an eighth aspect the present invention consists in a method of producing a desired polypeptide comprising the following steps:-

- 10 1. Transforming a recombinant strain of Dictyostelium with a recombinant plasmid vector including a DNA sequence encoding the desired polypeptide and sequences enabling the expression of the DNA sequence encoding the desired polypeptide;
- 15 2. Culturing the recombinant strain of Dictyostelium under conditions which allow the expression of the DNA sequence encoding the desired polypeptide and allowing the desired polypeptide to be produced either as a cell bound form or be secreted; and
- 20 3. Recovering the secreted desired polypeptide; characterised in that the recombinant plasmid vector includes an origin of replication derived from plasmid Ddp2 but lacks the functional genes for extrachromosomal replication in wild type Dictyostelium; and
- 25 that the recombinant strain of Dictyostelium includes a gene encoding a polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid.
- 30 As used herein the phrase "cell bound form" is intended to cover proteins either internal to the cell or present on the cell membrane.
- 35 In a preferred embodiment of this aspect of the present invention the gene encoding the polypeptide which facilitates the extrachromosomal replication of the

recombinant plasmid is present in a chromosome of the recombinant strain. Alternatively the gene is carried on a second recombinant plasmid present in the recombinant strain.

5 In a ninth aspect the present invention consists in a DNA molecule which includes a nucleotide sequence which encodes a polypeptide and which is capable of transforming Dictyostelium strains such that recombinant plasmid vectors which include an origin of replication derived 10 from a Ddp2-like plasmid, preferably plasmid Ddp2, are incapable of extrachromosomal replication in wild type Dictyostelium spp. are capable of extrachromosomal replication in the transformed Dictyostelium strain.

15 In a preferred embodiment of this aspect of the present invention the DNA molecule includes a sequence substantially as shown in Fig. 1 from nucleotide 2378 to nucleotide 5038, or part thereof.

As stated above, the present invention relates to the construction of extrachromosomal plasmid vectors for 20 Dictyostelium using much smaller sections of the plasmid Ddp2 than has previously been possible. The present invention enables the construction of plasmid vectors containing an origin of replication derived from Ddp2 which can be encoded on a section of Ddp2 DNA of less than 25 3.0 Kb, but omit sections of Ddp2 DNA that contain genes for polypeptides essential for replication and preferably DNA sequences that are unstable when cloned in E. coli. The replication of such plasmids can be achieved by maintaining them in recombinant strains of Dictyostelium 30 where the polypeptides required for plasmid replication are provided by genes inserted into the chromosomal DNA of the host cell or alternatively into another compatible plasmid vector. The present invention enables the production of a wide range of plasmid vectors which may be 35 constructed using the techniques known in the art and

- 13 -

disclosed herein, including plasmids designed for the expression of recombinant protein products in Dictyostelium spp.

The present invention further comprises the use of 5 these recombinant Dictyostelium strains for the maintenance of recombinant plasmids containing an origin of replication derived from Ddp2 but lacking functional genes for replication proteins. The maintenance of recombinant plasmids in hosts that have been genetically 10 modified to supply polypeptides necessary for plasmid replication is likely to be a crucial factor in the production of recombinant proteins using Dictyostelium spp.

SHORT DESCRIPTION OF THE DRAWINGS

In order that the nature of the present invention may 15 be more clearly understood preferred forms thereof will now be described with reference to the following examples and accompanying figures, in which:-

Figure 1 is the nucleotide sequence of the Dictyostelium plasmid Ddp2. The sequence of one strand of 20 DNA is shown, numbered clockwise from the SalI restriction enzyme site. The position of the recognition sites of restriction enzymes SalI, HindIII, BglII, NdeI, C1aI, EcoRI, EcoRV, PstI, BclI, XbaI, XhoI, AccI, HindII and ScaI are indicated. START and STOP indicates the position 25 of the first and last codons of the Rep gene respective.

KEY: A =Adenine. C =Cytosine. G =Guanine. T =Thymine;

Figure 2 is the amino acid sequence of the 30 polypeptide encoded by the Rep gene as derived from the DNA sequence of plasmid Ddp2. The nucleotide sequence of the coding strand of the Rep gene, numbered clockwise from the cleavage site of the SalI restriction enzyme, is aligned with the amino acid sequence predicted from the standard genetic code.

KEY: A =Adenine. C =Cytosine. G =Guanine. T =Thymine.

35 a =Alanine. c =Cysteine. d =Aspartic acid.

e =Glutamic acid. f =Phenylalanine.  
g =Glycine. h =Histidine. i =Isoleucine.  
k =Lysine. l =Leucine. m =Methionine.  
n =Asparagine. p =Proline.  
5 q =Asparagine. r =Arginine. s =Serine.  
t = Threonine. v =Valine. w =Tryptophan;

Figure 3 is a schematic representation of the major structural features of Ddp2 aligned with a map of the cleavage sites of some restriction enzymes;

10 Figure 4 is a schematic representation of the construction of plasmid pMUW111;

Figure 5 is a schematic representation of the construction of plasmid pMUW110;

15 Figure 6 is a schematic representation of the construction of plasmid pMUW102;

Figure 7 is a schematic representation of the construction of plasmid pMUW130;

20 Figure 8 is a schematic representation which summarizes the Ddp2 sequences used to construct plasmids pMUW111, pMUW102, pMUW110 and pMUW130;

Figure 9 is a schematic representation of the construction of the shuttle vectors pMUW1530 and pMUW1580;

25 Figure 10 is the nucleotide sequence of the shuttle vector pMUW1530. The sequence of one strand of DNA is shown, numbered anti-clockwise from the ClaI restriction enzyme site. The position of the recognition sites of restriction enzymes ClaI, ScaI, BamHI, BglII and NdeI are indicated.

KEY: A =Adenine. C =Cytosine. G =Guanine. T =Thymine;

30 Figure 11 is a schematic representation of the construction of the promoter and secretion signal sequence sections of an expression cassette in plasmid pMUW1594;

Figure 12 is a schematic representation of the cloning of the polyadenylation sequence from the  
35 Dictyostelium Actin 15 gene into plasmid pMUW1560;

- 15 -

Figure 13 is a schematic representation of the construction of the expression cassette in pMUW1621;

Figure 14 is a schematic representation of the construction of an expression vectors pMUW1630 and

5 pMUW1633 by insertion of the expression cassette into the shuttle vector pMUW1580; and

Figure 15 is the nucleotide sequence of the expression vector pMUW1630. The sequence of one strand of DNA is shown, numbered anti-clockwise from the *Clal* 10 restriction enzyme site. The position of the recognition sites of restriction enzymes *Clal*, *Scal*, *NsII*, *HindIII*, *SmaI* and *KpnI* are indicated. *START* indicates the position of the first codon of secretion signal peptide in the expression cassette.

15 KEY: A =Adenine. C =Cytosine. G =Guanine. T =Thymine.

Best Mode of Carrying Out the Invention

The present inventors have established for the first time the full nucleotide sequence of the Dictyostelium plasmid *Ddp2* as shown in Figure 1. The nucleotide 20 sequence has been numbered clockwise around the circular DNA molecule starting at the single cut site of the *SalI* restriction enzyme. Detailed examination of the DNA sequence of *Ddp2* has allowed different functional regions of the plasmid to be distinguished, as shown in Figure 3, 25 and regions likely to be unstable when cloned in *E. coli*. The elucidation of these different functional regions has allowed the present inventors to overcome a number of the technical problems that have hitherto limited the use of extrachromosomal vectors in Dictyostelium.

30 The DNA sequence of *Ddp2* between nucleotide 2378 and 5038 encodes a gene referred to herein as *Rep*. This section of *Ddp2* contains a large "open reading frame" where one of the six possible ways to read the triple nucleotide genetic code (known as codons) has a long 35 region without any of the codons that act as stop signals

for protein translation. Such an "open reading frame" considered along with flanking sequences that are similar to the promoter and poly-adenylation signals of previously described Dictyostelium genes (Kimmel & Firtel, 1982 In

5 The Development of Dictyostelium discoideum, Academic Press, New York, pp234-324) is strong evidence that the Rep gene could be transcribed into RNA and translated into a polypeptide containing 887 amino acids with the sequence shown in Figure 2. Evidence supporting the view that the

10 Rep gene is translated into a polypeptide comes from the inability of plasmids constructed with interruptions to the Rep gene, for example pMUW102, to replicate in wild type strains of Dictyostelium discoideum. The RNA and polypeptide product of the Rep gene has not yet been

15 detected and it is believed to be produced in only low amounts to positively regulate the initiation of plasmid replication by the host enzymes that normally replicate chromosomal DNA. However, it should be appreciated that either the messenger RNA or the translated polypeptide

20 derived from the Rep gene could be processed by the cellular biochemical machinery to produce one or more shorter polypeptides. It is also likely that the polypeptide also contains regions that act as negative regulators of plasmid copy number. None of these areas of

25 uncertainty subtract from the basic discovery that at least part of the open reading frame encodes a polypeptide that is essential for the replication of Ddp2. This finding explains the previously established need for shuttle vectors to contain a large section of Ddp2 DNA

30 since such vectors would need to contain both the origin of replication and an additional 2.66 kilobase pair Rep gene plus flanking control sequences.

Plasmid vectors based on Ddp2 need to contain DNA from the section of Ddp2 between the HindIII restriction enzyme site at 1153 base pairs and the BgIII restriction

enzyme site at 1885 base pairs.

This is demonstrated by the inability of plasmids that lack this section of DNA, for example pMUW110 (Figure 5), to replicate in wild type strains of 5 Dictyostelium. Plasmid pMUW110 contains the complete Rep gene plus flanking sequences including the polyadenylation sequences and 483 nucleotides encompassing the promoter region. Thus pMUW110 contains the sequences required to produce the polypeptide required for replication, but 10 lacks a functional origin of replication. Consequently, a Ddp2 origin of DNA replication or associated control sequences must lie before the BgIII restriction enzyme site at 1885 base pairs. This region of Ddp2 is present 15 in plasmid pMUW102 which contains the section of Ddp2 between the HindIII restriction enzyme site at 1153 base pairs and the XhoI restriction enzyme site at 3242 base pairs using plasmid pMUW102 (figure 6), but plasmid pMUW102 lacks a functional Rep gene and so is unable to replicate in wild type strains of Dictyostelium. The 20 presence of a functional origin of replication in plasmid pMUW102 is demonstrated by transforming it into Dictyostelium strains along with plasmid pMUW110 to provide the essential replication polypeptide from the Ddp2 Rep gene. The present inventors experimental results 25 clearly show that plasmid pMUW110 is inserted into the chromosomal DNA to form a stable recombinant strain of Dictyostelium and, in the same cells, plasmid pMUW102 is stably maintained as an extrachromosomal plasmid. This demonstration of an extrachromosomal plasmid containing an 30 origin of replication from plasmid Ddp2 and its maintenance in a Dictyostelium strain by virtue of chromosomal DNA containing the Rep gene encoding polypeptides essential for plasmid replication represents a significant technical advance. It is apparent to one 35 skilled in the art that similar techniques can be utilised

for the construction of a diverse range of plasmid vectors for Dictyostelium.

It is relevant to briefly examine the mechanism for selecting cells that were successfully transformed with 5 both pMUW102 and pMUW110. Both these vectors contain a selectable marker conferring resistance to the antibiotic G418, but other genes could be used to serve the same function. In fact the present inventors have developed another resistance gene bleomycin for use as a selectable 10 marker in Dictyostelium. The G418 resistance gene is under the control of Dictyostelium actin 6 promoter and the actin 8 3' poly-adenylation signals to ensure that it is expressed in Dictyostelium cells to provide a method of selecting the few cells that take up the plasmid DNA.

15 Plasmid pMUW110 which lacks an origin of replication can only be retained in those few cells where the plasmid becomes integrated into the chromosomal DNA. Any cells that are transformed with only plasmid pMUW102 can only be 20 resistant to G418 if the plasmid becomes integrated into the chromosomal DNA since this plasmid cannot replicate without the polypeptide produced by the Rep gene. However, some of the cells that receive both plasmids can have the plasmid pMUW110 integrated into the chromosomal 25 DNA in a manner that preserves the function of the Rep gene and so will be able to maintain multiple extrachromosomal copies of the plasmid pMUW102. Once the cells transformed with both plasmids pMUW102 and pMUW110 have been selected by resistance to G418 they may be 30 stably maintained in the absence of the antibiotic. Plasmid pMUW102 contains 2089 base pairs of Ddp2; a considerably smaller section of Ddp2 than previously known to be capable of extrachromosomal replication. This sequence has been substantially shortened by removing more of the Ddp2 DNA sequences that are not essential for the 35 replication of plasmid pMUW102 in recombinant strains of

5 Dictyostelium. The results with plasmid pMUW130 confirms that all the DNA sequences necessary for stable extrachromosomal replication at high copy number are contained in a 622 base pair HindIII-ClaI fragment of Ddp2. In the light of present knowledge as disclosed herein, it is also relatively simple to ascertain the essential sequences within the section of Ddp2 between the HindIII restriction enzyme site at 1153 base pairs and the ClaI restriction enzyme site at 1885 base pairs using 10 standard molecular biology techniques such as deletions and insertions. Experiments to determine the minimum section of Ddp2 DNA sequence necessary for plasmid vector construction have been carried out. Several copies of a TGTCATGACA sequence are essential for the function of the 15 Ddp2 origin of replication.

20 The use of smaller sections of Ddp2 for vector construction than previously possible allows the omission of some of the sequences likely to be responsible for plasmid instability in E. coli. Plasmid pMUW130 contains only one copy of sequences in the 501 base pair inverted repeat of Ddp2 and does not contain the long stretches of poly-adenine or poly-thymidine found between the end of the open reading frame and the SalI restriction enzyme site. Such inverted repeats and poly-adenine or 25 poly-thymidine sequences are known to be unstable in E. coli. Plasmid pMUW130 also omits the (GATGAA)11 repeat found at the end of the Rep gene and which is also likely to be unstable in E. coli. Therefore, it appears that the smaller sections of Ddp2 used to construct plasmid vectors 30 according to this invention have less of the problems of stability in E. coli than were previously encountered using larger segments of Ddp2 DNA.

35 The integrating plasmid pMUW110 contains all the information necessary for the controlled expression of the Ddp2 Rep gene required to maintain the copy number of

- 20 -

plasmid pMUW102. This control of plasmid copy number could not be predicted since there would be no direct linkage between the number of copies of the plasmid and the Rep gene as in the original plasmid. It is thought that this copy number control is probably achieved by an auto-regulatory mechanism where the product of the Rep gene represses further transcription from the Rep gene and so maintains a constant cellular concentration of the polypeptide that regulates plasmid replication. The localisation of the promoter sequences to the section of Ddp2 DNA between the BgIII restriction enzyme site and the start of the Rep gene, as disclosed herein, allows future experiments to determine the regulatory mechanisms governing the transcription of the open reading frame and control of plasmid copy number. It is anticipated that this approach will lead to experimental control of plasmid replication and copy number by suitable modification or duplication of the control sequences.

In the experiments described herein, the plasmid pMUW110 has been stably integrated into the Dictyostelium chromosomal DNA using the same selective marker, G418 resistance, as present on the extrachromosomal plasmid pMUW102. However, there would be advantages in using a different selective marker on the integrating vector from that used for the extrachromosomal plasmid. The present inventors have developed a thymidylate synthase gene as a second marker for selection in a Dictyostelium discoideum strain that is unable to synthesise thymidine (Chang et al, 1989, Nucleic Acids Research 17, 3655-3661). The thymidylate synthase selection has the advantage for biotechnological uses in that the selection is maintained in the absence of any antibiotic. Clearly any combination of selectable markers can be used on the integrating or extrachromosomal vectors, but the preferred combination is to have the thymidylate synthase marker on the

extrachromosomal plasmid and maintain it in the enzyme deficient Dictyostelium strain. This means that, without using any antibiotic selection, any host cell losing either the extrachromosomal plasmid or the functional 5 integrated vector would be unable to grow since any cell losing the production of the polypeptide necessary for plasmid replication would also lose the functions encoded on the extrachromosomal plasmid.

Examples of the application of the invention have 10 been demonstrated by the construction of a range of shuttle vectors and the production of a recombinant protein in Dictyostelium discoideum. The novel shuttle vectors pMUW1530, pMUW1570 and pMUW1580 incorporate the Ddp2 origin of replication on the 600 bp XbaI - ClaI 15 fragment (1175 - 1775 bp) of Ddp2 into a small E. coli plasmid (pMUW1510) that contains close to the minimal amount of sequence from pBR322 required for replication in E. coli in order to reduce the potential for these sequences to adversely effect the function of the shuttle 20 vector in D. discoideum. Other useful features of these shuttle vectors is that they contain very few sites for six base restriction enzymes, apart from single BamHI and ClaI sites in appropriate positions for the insertion of additional DNA without disrupting essential functions.

25 Sequences that might be inserted into such sites include genes for the production of recombinant proteins or selective markers, promoter sequences to control gene function and signal sequences for the correct processing of messenger RNA molecules and the translated proteins.

30 This is illustrated by the construction of a novel "expression cassette" suitable for the production and secretion of a recombinant proteins from Dictyostelium cells. This expression cassette contains the promoter from the D. discoideum actin 15 gene, a section of the D19 35 gene encoding a secretion signal peptide, the polylinker

from the E. coli plasmid pGEM3Z (for insertion of genes for expression) and lastly the polyadenylation signal from the D. discoideum actin 15 gene. However, it will be apparent to one skilled in the art that a wide range of 5 similar constructs could be made for this purpose using DNA sequences from other genes or even completely synthetic sequences serving the same functions.

The applications of the shuttle vector based on the technology disclosed in this document was demonstrated by 10 the production of a recombinant protein from an E. coli gene for enzyme B-glucuronidase from D. discoideum cells containing an expression vector constructed by inserting the expression cassette into the shuttle vector pMUW1580.

Plasmid Ddp2 is believed to be the first functionally 15 characterized member of a new group of structurally and functionally similar plasmids. This new group of plasmids can be defined as all encoding a single polypeptide of 700-1000 amino acids which is essential for plasmid replication and which has sequence homologies with the 20 Ddp2 Rep gene, indicating a common evolutionary origin. Further, the origin of replication of these plasmids is associated with one arm of an inverted repeat sequence that is distinct from the Rep gene. The inventors confidently predict that the techniques they have 25 disclosed in this application can be used to construct further extrachromosomal plasmid vectors for use in the biotechnology industry starting from the functionally analogous regions of any of this broader group of "Ddp2 - like" plasmids.

30 The only other member of this "Ddp2-like" group of plasmids to have been sequenced to date is plasmid pDG1 isolated from a unidentified Dictyostelium species (Orii et al (1987) Nucleic Acids Res. 15,1097-1107). Plasmid pDG1 has a very similar structure to Ddp2, possessing 35 similar sized inverted repeats and a single open reading

frame analogous to the Rep gene of Ddp2. Despite plasmid pDG1 having been fully sequenced, nothing is known regarding the functions of these features or the location of the origin of replication (Orii et al (1989) Nucleic Acids Res. 17,1395-1408). The only recombinant shuttle vector produced with pDG1 sequences incorporated the long, 4.2 Kb Clal fragment of pDG1, i.e., omitting only 0.2 Kb from the whole plasmid (Orii et al (1989) Nucleic Acids Res. 17,1395-1408). Such pDG1 based plasmids are very unstable in E. coli (Saing et al (1988) Mol. Gen. Genet. 214,1-5) and so are unsuitable for use in the production of recombinant proteins.

The plasmid pDG1 is recognized as a member of the "Ddp2-like" group of plasmids by virtue of its having a similar structure and having sequence homologies with Ddp2 in the region of the open reading frame at both the DNA and amino acid levels. The non-coding regions of these two plasmids have little sequence homology, apparently being free to diverge in the course of evolution. The presence of large inverted repeats in both pDG1 and Ddp2 is probably not a key feature of the group of "Ddp2-like" plasmids as only one copy is essential for the replication of Ddp2.

In the light of the functional data from the analogous regions of Ddp2, as disclosed in this application it is possible to re-evaluate the pDG1 sequence data and predict that pDG1 origin of replication lies outside the operating reading frame and overlaps with one of the inverted repeats. In addition, the speculation (Orii et al (1989) Nucleic Acids Res. 17,1395-1408) concerning the weak homologies of the Rep gene with reverse transcriptase is unlikely to be correct as the homology is not conserved in Ddp2. The Rep gene of Ddp2 can be aligned with the open reading frame of pDG1 with 35% of amino acids in identical positions indicating

considerable evolutionary homologies. The proteins encoded by the two plasmids also have similar structures, being comprised of two similar sized domains separated by a threonine rich sequence and the carboxy terminus of both 5 proteins being a highly acidic glutamic and aspartic acid rich sequence. To one skilled in the art, the similarities between the proteins produced by these two plasmids indicates they have very similar functions and also indicates regions of high sequence homology which are 10 most likely to have roles crucial for the proteins function. Whilst it is unlikely that the protein from pDG1 would be sufficient to cause replication of the Ddp2 origin of replication (and vice versa) because the sequence recognized by the protein is likely to be 15 specific to the individual origin of replication, it is very likely that novel proteins constructed from sections of both proteins would function correctly. For example, the replacement of the acidic carboxy terminus of the Ddp2 Rep protein with the carboxy terminus of the pDG1 protein 20 should not affect the ability of the molecule to allow replication from the Ddp2 origin of replication. Furthermore, it should be possible to change the specificity of the Ddp2 Rep gene simply by replacing the section of the protein that recognizes the Ddp2 origin of 25 replication by a section recognizing an origin of replication from another member of the "Ddp2-like" group of plasmids. Clearly, the basic technology disclosed in this application, whereby, the replication protein and the origin of replication are separated onto separate vectors, 30 is capable of a wide range of different applications for the construction of plasmid vectors incorporating sections from the broad group of "Ddp2-like" plasmids.

Example 1

Sequencing of plasmid Ddp2

35           Our laboratory at Macquarie University sequenced Ddp2

by cutting Ddp2 DNA into many small fragments and cloning them separately into a commercially available plasmid called pGEM3Z (Promega Corporation, Madison, USA). In this vector, small sections of Ddp2 DNA were stable and 5 could be sequenced using a technique called "double stranded sequencing" where a small oligonucleotide is used to prime the synthesis of a new radio-labelled DNA strand on a template of denatured plasmid DNA. The oligonucleotide primer can be the complementary sequence 10 to the SP6 or T7 regions flanking the cloning site or it can be a custom synthesised oligonucleotide with a sequence that matches part of the cloned Ddp2 DNA.

Ddp2 DNA was digested with the restriction enzymes C1aI, Sau3A, AluI or RsaI and cloned into the plasmid 15 pGEM3Z at the AccI, BamHI or SmaI restriction enzyme sites using standard molecular biology techniques, and transformed into the E. coli strain JM109. Clones containing Ddp2 DNA were selected at random and stored in broth containing 15% glycerol and stored at -80 degrees.

20 Plasmid DNA from the clones was prepared using alkaline lysis and a RNase enzyme treatment as recommended by the Promega literature on pGEM3Z. Before use in the sequencing reaction, 4ug of each plasmid was alkaline denatured with a brief treatment with 0.4M sodium 25 hydroxide, precipitated with ethanol and annealed with 10 picomoles of oligonucleotide primer according to the procedure recommended by Pharmacia LKB Biotechnology (Uppsala, Sweden) for their T7 DNA polymerase sequencing kit which was used for the sequencing reaction. The 30 sequencing reaction used ATP radio-labelled with  $^{35}\text{S}$ . The radio-labelled DNA was separated on 6% acrylamide/8M urea gels which were then fixed in 10% methanol plus 10% acetic acid, dried and autoradiographed. The sequence 35 revealed by the autoradiography films were entered into a computer and then overlapping sequences matched.

- 26 -

automatically and compiled into the complete DNA sequence of Ddp2.

The full sequence of Ddp2 is available from the EMBL data base, accession number X51478.

5 Example 2

Location of the Origin of Replication of Ddp2

In further experiments the Ddp2 origin of replication was located to within the HindIII - ClaI fragment (1153-1775 bp) of Ddp2 as in plasmid pMUW130.

10 pMUW111

The plasmid pMUW111 was constructed by inserting the 4.1 Kb HindIII to ScaI fragment of Ddp2 into the SalI site of BIOSX. BIOSX is an integrating D. discoideum/E. coli shuttle vector constructed by Nellen *et al.* (Gene. 39 15 (1985) 155-163) and contains the Ampicillin and Kanamycin/G418 antibiotic resistance genes.

Ddp2 plasmid was first digested with restriction enzymes HindIII and ScaI. After the digestion was completed, the Hind III 5' overhang ends were made blunt 20 using an end-filling reaction involving the enzyme DNA polymerase I "Klenow fragment". After this reaction was completed, it was fractionated in a 0.8% TBE agarose gel. The 4.1 Kb fragment was then excised from the gel and purified using a commercial kit, "Gene-Clean" 25 (BIO101, Inc., USA). The purified DNA was then ligated with BIOSX that had been digested with SalI and end-filled. After ligation, the mixture was transformed into E. coli strain CES201 (Leach, D.R.F. and Stahl, F.W. (1983). Nature 305, 448-451). CES201 was made competent 30 for transformation using the procedure as published by Hanahan, D. (J. Mol. Biol. (1983) 166, 557-580). The transformation mixture was then plated onto Luria-agar containing 50ug/ml ampicillin. E. coli ampicillin resistance transformants containing pMUW111 were confirmed 35 by restriction fragment mapping of isolated plasmids and

- 27 -

also by radioactive hybridization using Ddp2 as a probe.

10 ug of pMUW111 was then used to transform Dictyostelium axenic strain, AX3K, using the standard calcium phosphate precipitation procedure developed by 5 Nellen W. et al. (Mol. Cell. Biol. (1984) 4, 2890-2898) with G418 selection. To determine if pMUW111 was capable of autonomous replication, total nuclear DNA was isolated from G418 resistant transformants and then screened on a "lysis in the gel" as described by Noegel A. et al 10 (J. Mol. Biol. (1985) 185, 447-450). The gel was then southern-transferred onto Zeta-probe blotting membrane (Bio-RAD) and hybridized using <sup>32</sup>P-labelled Ddp2 DNA. Autoradiography showed that pMUW111 had a higher mobility than the bulk chromosomal DNA, indicating it existed as an 15 autonomously replicating plasmid.

pMUW102

The plasmid pMUW102 was constructed by inserting the 3.2 Kb SalI to XhoI fragment of Ddp2 into the Sal I site of BIOSX. This fragment contained only part of the open 20 reading frame. Hence a complete functional protein(s) would not be expected to be produced by this construct.

Ddp2 plasmid was first digested with restriction enzymes SalI and XhoI. The sample was then fractionated in a 0.8% TBE agarose gel. The 3.2 Kb fragment was then 25 excised from the gel and purified using a commercial kit, "Gene-Clean". The purified DNA was then ligated with BIOSX that had been digested with SalI. After ligation, the mixture was transformed into competent E. coli strain CES201. The transformation mixture was then plated onto 30 Luria-agar containing 50 ug/ml ampicillin. E. coli ampicillin resistant transformants containing pMUW102 were confirmed by restriction fragment mapping of isolated plasmids and also by radioactive hybridization using Ddp2 as a probe.

35 10ug of pMUW102 was then used to transform

D. discoideum axenic strain, AX3K, using standard calcium phosphate precipitation procedure with G418 selection. To determine the fate of pMUW102, total nuclear DNA was isolated from G418 resistant transformants and then 5 screened on a "lysis in the gel". The gel was then southern-blotted onto Zeta-probe blotting membrane and hybridized using <sup>32</sup>P-labelled Ddp2 DNA. Autoradiography showed that pMUW102 had the same mobility as the bulk chromosomal DNA, indicating it had integrated into 10 chromosomal DNA and it was not capable of existing as a free plasmid. This experiment demonstrated that an intact open reading frame is essential for existence as an autonomously replicating plasmid.

pMUW110

15 The plasmid pMUW110 was constructed by inserting the 3.4 Kb BglII to ScaI fragment of Ddp2 into the Sal I site of BIOSX. This fragment contained the whole open reading frame "Rep gene" and the 5' and 3' flanking sequences that control the production of protein(s) specified by the open 20 reading frame.

Ddp2 plasmid was first digested with restriction enzymes ScaI and BglII. After the digestion was completed, the BglII 5' overhang ends were made blunt using an end-filling reaction involving the enzyme DNA 25 polymerase I "Klenow fragment". After this reaction was completed, the sample was fractionated in a 0.8% TBE agarose gel. The 3.4 Kb fragment was then excised from the gel and purified using a commercial kit, "Gene-Clean". The purified DNA was then ligated with 30 BIOSX that had been digested with SalI and end-filled.

After ligation, the mixture was transformed into E. coli strain CES201 that had been made competent for transformation. The transformation mixture was then plated onto Luria-agar containing 50 ug/ml ampicillin. 35 E. coli ampicillin resistant transformants containing

pMUW110 were confirmed by restriction fragment mapping of isolated plasmids and also by radioactive hybridization using Ddp2 as a probe.

10ug of pMUW110 was then used to transform

5 D.discoideum axenic strain, AX3K, using standard calcium phosphate precipitation procedure with G418 selection. To determine the fate of pMUW110, total nuclear DNA was isolated from G418 resistant transformants and then screened on a "lysis in the gel". The gel was then

10 southern-transferred onto Zeta-probe blotting membrane and hybridized using <sup>32</sup>P-labelled Ddp2 DNA. Autoradiography showed that pMUW110 had the same mobility as the bulk chromosomal DNA, indicating it had integrated into the chromosomal DNA and it was not capable of existing as a

15 free plasmid.

The difference between pMUW111 and pMUW110 is that 732 nucleotides between the HindIII restriction enzyme site at 1153 base pairs and the BglII restriction enzyme site at 1885 base pairs is missing in pMUW110. Hence the

20 inability of pMUW110 to exist as a plasmid in AX3K could be explained by one of the following:

i) The 732bp sequence contained part of the origin of replication (ORI) of the plasmid Ddp2.

ii) The 732bp sequence contained cis acting element(s) that control the production of protein(s) specified by the

25 open reading frame.

The first explanation was found to be correct by a subsequent experiment involving the co-transformation of AX3K with both pMUW102 and pMUW110. Screening of the

30 G418-resistant transformants revealed that pMUW102 had a higher mobility than the bulk chromosomal DNA. This proved that pMUW102 could exist as an extrachromosomal plasmid only in the presence of pMUW110, which contained the intact open reading frame and hence is capable of

35 providing the transacting protein(s) required for pMUW102

to replicate as a plasmid.

pMUW130

The plasmid pMUW130 was constructed by inserting the 622 base pair HindIII to ClaI fragment from Ddp2 (ie 1153 5 base pair to 1775 base pair) into the commercial E. coli plasmid pGEM3Z (Promega Corporation, Madison, USA) which had been digested with AccI and HindIII restriction enzymes. The construction of the plasmid used the same procedure as that of pMUW102 (above) except that the 10 E. coli strain used was HB101.

Plasmid pMUW130 contains most of the 732 base pairs sequence that are in plasmid pMUW102, but not in plasmid pMUW110 and which was thought to be required for extrachromosomal replication. An experiment where pMUW102 15 and pMUW110 were co-transformed into D. discoideum strain AX3K demonstrated that pMUW130 can replicate extrachromosomally in the presence of pMUW110 which has been integrated into the chromosomal DNA. This confirms that an origin of DNA replication is located on this small 20 HindIII - ClaI fragment of Ddp2 DNA. At approximately 3.3 kilobase pairs of DNA, pMUW130 was substantially smaller than previous shuttle vectors that had been constructed for Dictyostelium spp.

The location of an origin of replication on the 25 HindIII - ClaI fragment incorporated into plasmid pMUW130 raises interesting scientific questions as to whether the similar sequences that occur in the small HindIII fragment (66-1153 bp) are also capable of acting as an origin of replication. This was investigated by cloning the small 30 HindIII fragment (66-1153 bp) into the Hind III site of plasmid B10SX to form plasmid pMUW105. However, plasmid pMUW105 was unable to replicate extrachromosomally when mixed with plasmid pMUW110 (to provide the Rep gene) and transformed into D. discoideum strain AX3K. The small 35 HindIII fragment in pMUW105 contains an entire, near

- 31 -

perfect copy of the 501 bp inverted repeat sequence that forms most of the Ddp2 origin of replication in plasmid pMUW130. So the failure of pMUW105 to replicate extrachromosomally demonstrates that either the sequences 5 just outside the 501 bp inverted repeat are essential for replication or the 11 nucleotide substitutions between the two copies of the 501 bp inverted repeat have prevented the copy in the small HindIII fragment in pMUW105 from acting as the origin of replication. Both of these 10 possibilities result in the absence of or changes to copies of the DNA sequence

TTTTTGTCACTTTTTTTGTCACTTGTCA, one copy of which lies just outside the 501 bp inverted repeat in pMUW130 and while a second copy of which is altered in pMUW105.

15 This sequence contains two copies of a 10 bp palindrome TGTCACTTGTCA (i.e. the two halves are symmetrical, so the complementary DNA strand will have the same sequence in the opposite orientation). Such palindromic sequences are typical of many sites recognized by DNA binding proteins, 20 which would be consistent with this sequence being important for regulation of the origin of replication.

The Ddp2 origin of replication in plasmid pMUW130 contains two copies of the above oligo T sequence, each of which contains two palindromes. Deletion of one copy of 25 the sequence by cutting out the HindIII - BglII restriction fragment (1153-1369 bp, numbered according to Ddp2) of plasmid pMUW130 produced plasmid pMUW138 which is unable to replicate extrachromosomally in D. discoideum, thus demonstrating the importance of this sequence for the 30 function of the origin of replication. However, it is unlikely that this sequence is the actual origin of replication, which is believed to lie in flanking sequences.

Example 3Construction of a Small Shuttle Vector

A list of oligonucleotide sequences used in vector constructions is shown in Table I.

5 Despite plasmid pMUW130 being a great improvement on all shuttle vectors previously available for D. discoideum, it has some drawbacks for use in the biotechnology industry. Plasmid pMUW130 contains a disrupted polylinker (concentrated region of restriction 10 enzyme sites) and DNA sequences derived from the Lac operon and the parent pBR322 plasmid which are not required in a Dictyostelium vector.

Ideally, the restriction enzyme sites in an expression plasmid should be only in positions convenient 15 for the manipulation of the gene to be expressed and the amount of unnecessary DNA should be minimized. Shuttle plasmid pMUW 1530 was designed specifically for the purpose of easy manipulation of inserted sequences. This plasmid contains the minimal sequences derived from pBR322 20 that allow replication in E. coli plus the ampicillin resistance selective marker. The "poison sequences" that are known to interfere with replication from the SV40 origin of replication (Lusky & Botchan (1981) Nature 293, 79-81.) and gene expression in mammalian cells (Peterson 25 et al (1987) Mol. Cell. Biol. 7,1563-1567) were excluded, although as yet their influence on D. discoideum plasmids is unknown. Other features of the plasmid include the creation of two unique six base restriction sites (BamHI and ClaI) positions suitable for the insertion of 30 expression cassettes or selective markers.

Table 1.LIST OF OLIGONUCLEOTIDE SEQUENCES USED IN  
VECTOR CONSTRUCTION.

5 The sequence (5' to 3') of the oligonucleotides synthesised at Macquarie University is shown together with the approximate position of restriction enzyme cutting sites.

PCR primers for cloning the actin15 promotor

GA190. TGGCCAAGCTTAGATCTACAAATTAATTAATCCC  
EaeI HindIII BglII

10 GA188. CCCGGGATGTTACCATGCATTTTATTTTTA  
SmaI/AvaI FokI NsiI

PCR primers for cloning the actin15 3' region

GA189. TGCCGGTACCTAAATCATGAATGAAAGTGCT  
KpnI

15 GA186. CCCGGGAATTCAAGATCTTTCAATGGAGATTTGAT  
SmaI/AvaI EcoRI BglII

PCR primers for cloning the secretion signal from the D19 gene

GA187. GGGAAAGCTTGGATGAATTCAAAAAATGAAATTCCAACAT  
HindIII FokI EcoRI

20 GA182. CCCGGGTCGACCTGCTATTGCATTTGCATATGTTAA  
SmaI/AvaI SalI BspMI NdeI

Linker inserted into NdeI site to complete secretion signal sequence

GA297. TACGCCAATGCATATGAAAGCT  
NsiI HindIII  
25 NdeI

GA296. TAAGCTTCATATGCATTGGCG  
HindIII NdeI NsiI

PCR primers used to clone pGEM3Z origin of replication

GA181. GGGGTGGATCCGCTAGCCGCATCGATAGGTGGCACTTTCGG  
BamHI NheI ClaI

30 GA179. GGAGGGATCCAAAGGCCAGCAAAAGGCCAGCAAAAGGC  
BamHI

Sequencing oligonucleotide for pMUW1410

GA220 GAAGCATTATCAGGG

35 Linker used to clone the gene for B-glucuronidase

GA310 AATTCCCGGG  
EcoRI SmaI

pMUW1410

Plasmid pMUW1410 is an E. coli plasmid which was made to be the basis for construction of a series of shuttle vectors, including pMUW1530.

5 Plasmid pMUW1410 was constructed using two synthetic oligonucleotides GA179 and GA181 as primers to amplify the required pGEM3Z sequence in a polymerase chain reaction (PCR). The two oligonucleotide primers were each designed as two sections, the 5' end of the sequences containing  
10 restriction sites required for cloning and the 3' end of the sequences specifically matching the sequence of the plasmid pGEM3Z. The 3' ends of the oligonucleotide GA179 is the same as the pGEM3Z nucleotides 452-472 bp (Promega Corp. numbering system) while the 3' end of  
15 oligonucleotide GA181 is complementary to pGEM3Z nucleotides 2254-2240 bp, i.e. they prime opposite strands of the pGEM3Z DNA during the PCR reaction.

The PCR reaction was carried out using 10ng of pGEM3Z cut with restriction enzyme PvuII to linearized the  
20 plasmid, 20pico moles of each oligonucleotide, 0.03 mM of each of the four deoxynucleotide triphosphates dATP, dTTP, dCTP and dGTP, Taq polymerase buffer (Biores) to a final volume of 50 ul and 1.25 units of Taq polymerase (Biores). The reaction was carried out for eight cycles  
25 using 120 second incubations at 95 degrees to denature, 50 degrees to anneal and 72 degrees for the extension reaction. The polymerase was removed from product of the PCR reaction by extracting with phenol, then chloroform and the DNA precipitated with ethanol at -20 degrees.  
30 The product of the PCR (which consisted of the pGEM3Z sequence 452-2254 bp flanked by the sequences of the two oligonucleotides GA179 and GA181) was then digested with the restriction enzyme BamHI to cleave the BamHI sites at the 5' end of the two oligonucleotides, and then the  
35 enzyme removed by extraction with 50% phenol/chloroform,

chloroform and then the DNA was precipitated with three volumes of ethanol at -70 degrees. Finally, the DNA product of the PCR reaction was self ligated using the BamHI sticky ends to form intact plasmids and the plasmids 5 transformed into the E. coli strain Dh5a(Bethesda Research Laboratories) by electroporation using the procedures recommended by Biorad, the manufacturer of the "Gene pulser" equipment. The transformed cells were spread onto LB agar containing 100 ug ampicillin per ml. E. coli 10 clones resistant to ampicillin were selected, their plasmids (e.g. pMUW1410) prepared by alkaline lysis and checked for size and the desired pattern of restriction enzyme sites using agar electrophoresis.

The plasmid pMUW1410 was approximately 1.8 Kb in size 15 as expected for the desired portion of pGEM3Z (452-2254 bp) containing the pBR322 origin of replication and the ampicillin gene. Indeed, the ability of the E. coli clone containing pMUW1410 to replicate on ampicillin agar means the plasmid must contain a functional origin of 20 replication and the ampicillin resistance gene. pMUW1410 also contains restriction sites for ClaI, BamHI and NheI derived from the synthetic oligonucleotides. The sequence of the plasmid pMUW1410 in the region of the BamHI site was confirmed using a T7 polymerase sequencing kit 25 (Pharmacia) and a synthetic oligonucleotide GA220 which is designed to anneal to the ampicillin gene (2149-2164 bp, pGEM3Z numbering) so that the sequencing reaction covers the sequence derived from the oligonucleotides GA179 and GA181. The sequencing reaction confirmed that the 30 oligonucleotides GA179 and GA181 used to create pMUW1410 had in fact bound to the expected positions in pGEM3Z and excludes the possibility of errors due to miss-priming at any other position.

pMUW1530

Shuttle vector pMUW1530 was constructed by inserting the XbaI - ClaI fragment (1175-1775 bp) of Ddp2 containing the origin of replication into the NheI and 5 ClaI sites of plasmid pMUW1410.

Plasmid pMUW1015 containing the large AluI (1155-3223 bp) fragment of Ddp2 was used as the source of the Ddp2 origin of replication. 10 ug of pMUW1015 was digested with XbaI and EcoRI restriction enzymes and a 1.2 Kb DNA 10 fragment (i.e. 1175-2436 bp of Ddp2) isolated by agarose gel purification. The appropriate DNA band was excised from the electrophoresis gel and frozen to disrupt the gel matrix. The DNA was extracted using the centrifugation methods of Heery et al ((1990) TIG 6,173.) and then 15 phenol/chloroform extracted and ethanol precipitated to remove traces of the ethidium bromide stain. The DNA was further digested with the ClaI restriction enzyme and the 0.6 Kb XbaI - ClaI fragment (1175-1775 bp, Ddp2 numbering) gel purified as 20 described above.

Plasmid pMUW1410 was digested with the restriction enzyme NheI and subsequently with enzyme ClaI, since the NheI site is too close to the ClaI site to cut efficiently after the ClaI enzyme has cut. The digestion was then 25 dephosphorylated by adding 1/40th volume of 20% SDS, 1/6th volume of 1M Tris buffer pH 9.0 and then 1 unit of Calf intestinal alkaline phosphatase (Boehringer) and incubating at 37 degrees for one hour. The enzyme was then removed by extracting with 50% phenol/chloroform 30 followed by chloroform extraction and then the DNA precipitated with ammonium acetate and two volumes of ethanol.

The XbaI - ClaI fragment from plasmid pMUW1015 (i.e. the Ddp2 origin of replication) prepared above was ligated 35 into the plasmid pMUW1415 (cut with NheI and ClaI and

treated with alkaline phosphatase), transformed into the E. coli strain "Sure" (Statagene) and plated onto LB agar containing 100 ug ampicillin per ml. E. coli clones resistant to ampicillin were selected, their plasmids (e.g. pMUW1530) prepared by alkaline lysis and checked for size and the desired pattern of restriction enzyme sites using agar electrophoresis.

Plasmid pMUW1530 is a 2.4 Kb shuttle plasmid containing the Ddp2 origin of replication inserted into the NheI and ClaI sites of plasmid pMUW1410. Evidence confirming this includes the presence of the BglII and NdeI sites from the Ddp2 origin of replication at the expected distance from the BamHI and ClaI sites found in pMUW1410. pMUW1530 does not contain the XbaI or NheI restriction sites used for cloning since the compatible "sticky ends" were destroyed by the ligation.

5ug of pMUW1530 mixed with 5ug of plasmid pMUW110 was then used to transform D. discoideum axenic strain, AX3K, using the standard calcium phosphate precipitation procedure with G418 selection. G418 resistant transformants were screened by "lysis in a gel", southern blotting onto zeta-probe membrane and probed with <sup>32</sup>P labelled pGEM3Z. This demonstrated the presence of an extrachromosomal plasmid with the size of plasmid pMUW1530 containing pGEM3Z DNA sequences.

#### pMUW1570

Shuttle vector pMUW1570 is the same as pMUW1530, but with the NdeI restriction site removed to allow NdeI to be used for the manipulation of genes cloned into the plasmid.

Plasmid pMUW1530 was digested with the NdeI restriction enzyme in 11 ul of 10mM Tris buffer pH 7.5, 10mM MgCl<sub>2</sub> and 50mM NaCl. The ends of the DNA were then filled by simply adding 1 unit of T7 polymerase and 3ul of the "C long" mix of deoxynucleotides supplied with the Pharmacia T7 polymerase sequencing kit and incubating at

room temperature for five minutes. The plasmid was then religated by adding 2ul ligation buffer (Boehringer), adjusting the volume to 20ul by adding water and 1 unit of T4 ligase and then incubating at 4 degrees overnight. The 5 religated plasmid was transformed into the E. coli strain "Sure" (Statagene) and plated onto LB agar containing 100 ug ampicillin per ml. E. coli clones resistant to ampicillin were selected, their plasmids (e.g. pMUW1570) prepared by alkaline lysis and checked for size and the 10 absence of the NdeI restriction site.

pMUW1580

Shuttle vector as pMUW1580 is the same pMUW1570, but with the BglII restriction site removed to allow BglII to be used for the manipulation of genes cloned into the 15 plasmid.

Plasmid pMUW1530 was digested with the NdeI restriction enzyme, end filled with T7 polymerase, self ligated and transformed into E. coli using the same procedures as for pMUW1570. E. coli clones resistant to 20 ampicillin were selected, their plasmids (e.g. pMUW1580) prepared by alkaline lysis and checked for size and the absence of the BglII restriction site.

Plasmid pMUW1580 contains a second ClaI site created by end filling the BglII site. However, in most strains 25 of E. coli this sequence is methylated so that the ClaI enzyme will not cut the new ClaI site.

5ug of pMUW1580 was mixed with 5 ug of plasmid pMUW110 and used to transform the D. discoideum axenic strain, AX3K, using the standard calcium phosphate 30 precipitation procedure with G418 selection. G418 resistant transformants were screened by "lysis in a gel", southern blotting onto Zeta-probe membrane and probed with <sup>32</sup>P labelled pGEM3Z. This demonstrated an extrachromosomal plasmid with the size of plasmid pMUW1580 35 containing pGEM3Z DNA sequences. Thus, plasmid pMUW1580

is a small, 2.4 Kb shuttle vector containing the minimum number of six base restriction sites, which is particularly suitable for use in the construction of expression vectors.

5 Example 4

Construction of an Expression Cassette

An "expression cassette" is a single, easily cloned piece of DNA which contains in their correct relative positions all the sequences required to ensure expression 10 of a gene and the correct processing of the messenger RNA and protein product. Usually the cassette contains a number of restriction sites (polylinker) behind the promoter in a good position for inserting the gene to be expressed. The use of a well designed expression cassette 15 greatly facilitates the expression of a range of genes and is much preferred to the alternative of cloning all the necessary DNA sequences on an adhoc basis.

We have designed a novel expression cassette specifically for insertion into the BamHI site of the 20 shuttle vectors pMUW1530, pMUW1570 and pMUW1580. The expression cassette is designed to minimize the amount of unnecessary DNA sequences and restriction sites. This was achieved by cloning the required control and signal sequences using PCR techniques to insert at key positions 25 the restriction sites required for cloning, using sites that can be destroyed during the construction procedure. The cassette contains a promoter from the D. discoideum actin 15 gene, a sequence coding for a secretion signal peptide, a polylinker containing restriction sites 30 allowing the insertion of genes for expression and a polyadenylation signal sequence from the D. discoideum actin 15 gene.

Each component section of the expression cassette was cloned separately and then assembled into the complete 35 cassette inside the polylinker of pGEM3Z.

Cloning the Actin 15 Promoter, Plasmid pMUW1480

The actin 15 promoter was selected because it is well characterised and is known to be expressed at a relatively high level soon after the onset of starvation (Cohen et al 5 (1986) EMBO J. 5,3361-3366). For the purpose of the production of recombinant proteins, this pattern of expression is desirable to avoid the protein being produced during active growth where the resulting metabolic drain may cause a selective advantage for any 10 non-secreting mutants.

The two synthetic oligonucleotides GA190 and GA188 were used as primers to amplify the required actin 15 promoter sequence in a polymerase chain reaction (PCR). The two oligonucleotide primers were each designed as two 15 sections, the 5' end of the sequences containing restriction sites required for cloning and the 3' end of the sequences specifically matching the sequence of the Actin 15 gene in plasmid pTS1 (Chang et al (1989) Nucleic Acids Res. 17,3655-3661). The 3' ends of the 20 oligonucleotide GA190 is the same as the promoter nucleotides between -247 and -230 (numbering back from A of the ATG start codon) while the 3' end of oligonucleotide GA188 is complementary to nucleotides between +3 and -13, i.e. they prime opposite strands of 25 the actin 15 DNA during the PCR reaction.

The PCR reaction was carried out using 30 ng of pTS1 cut with restriction enzymes PvuII and ScaI to ensure the plasmid is unable to replicate during later cloning steps, 30 20p moles of each oligonucleotide, 0.03 mM of each of the four deoxynucleotide triphosphates dATP, dTTP, dCTP and dGTP, Taq polymerase buffer (Biores) to a final volume of 50 ul and 1.25 units of Taq polymerase (Biores). The reaction was carried out for ten cycles using 120 second incubations at 95 degrees to denature, 40 degrees to 35 anneal and 72 degrees for the extension reaction. At the

end of the PCR reaction, 1 unit of T4 polymerase was added and incubated at room temperature for 15 minutes to ensure the ends of the DNA were blunt. 20 ug of glycogen in 1ul (Boehringer) was added and 2 ul of acetate buffer (to aid 5 precipitation of the small DNA fragments) before the polymerases were removed by extracting with 50% phenol in choroform, then chloroform and the DNA precipitated with three volumes of ethanol at -70 degrees.

The product of the PCR reaction (consisting of the 10 Actin 15 promoter sequence between -247 and +3 relative to the start codon flanked by the sequences of the two oligonucleotides GA190 and GA188) was shown to have the expected size of approximately 300 bp by electrophoresis in 1.6% agarose against size markers (BRESA) of phage 15 SPP-1 digested with the restriction enzyme EcoRI.

The DNA product of the PCR reaction was mixed with 100ng of pGEM3Z which had been cut with the restriction enzyme SmaI to create blunt ends. The mixture was ligated with 3 units of T4 ligase in ligation buffer for two hours 20 at room temperature and then precipitated with ammonium acetate and two volumes of ethanol. The religated plasmids were transformed into the E. coli strain Dh5a (Bethesda Research Laboratories) by electroporation using the procedures recommended by Biorad, the manufacturer of 25 the "gene pulser" equipment. The transformed cells were plated onto LB agar containing 100 ug ampicillin per ml, 0.5mM IPTG (isopropyl-B-d-thiogalactopyranoside) and 50ug X-Gal (5-bromo-4-chloro-3-indolyl-B-galactoside) per ml. E. coli clones resistant to ampicillin and producing large 30 white colonies (indicating the plasmid has DNA inserted into the polylinker) were selected, their plasmids (e.g. pMUW1480) prepared by alkaline lysis and checked for size and the desired pattern of restriction enzyme sites using agar electrophoresis.

35 The plasmid pMUW1480 digested by the restriction

enzyme *Pvu*II produced a fragment with approximately of 700 bp, comprised of 379 bp of pGEM3Z sequences containing an approximately 300 bp insert, as expected for the desired actin 15 promoter (250 bp) flanked by the sequences 5 derived from the synthetic oligonucleotides GA190 and GA188. pMUW1480 also contains restriction sites for HindIII, BglII, NsiI and FokI derived from the synthetic oligonucleotides. The identity of the promoter inserted into plasmid pMUW1480 was confirmed by sequencing using a 10 T7 polymerase sequencing kit (Pharmacia) and commercially supplied oligonucleotides (Promega) which anneal to SP6 and T7 regions flanking the polylinker. The sequencing excludes any possibility of errors in the sequence.

Cloning a Sequence for a Secretion Signal, Plasmid pMUW1450

15 Secretion of a protein requires a signal sequence at the amino terminal end of the polypeptide. This signal peptide is the first part of the protein to be transcribed and causes the ribosome to bind to the endoplasmic reticulum membranes and feed the nascent polypeptide 20 across the membrane into the lumen of endoplasmic reticulum. Subsequently, the signal peptide is cleaved from the rest of the protein.

25 The *D. discoideum* protein PsA possesses a 20 amino acid signal peptide which has characteristics typical of many eukaryotic signal peptides (Perlman & Halvorson (1983) J. Mol. Bio. 167, 391-409) and so it should be a reliable signal to use of the secretion of recombinant proteins.

30 The two synthetic oligonucleotides GA187 and GA182 were used as primers in a PCR reaction to amplify the DNA sequence coding for the PsA signal peptide from the D19 gene encoding the PsA protein (Early et al (1988) Mol. Cell. Biol. 8, 3458-3466). The methods used were the same as described for cloning the actin 15 promoter (see 35 above). However, some difficulty occurred in cloning the

correct product of the PCR reaction.

The plasmid pMUW1450 gave the correct size fragment when digested by the restriction enzyme *Pvu*II, but when the insert was sequenced it was found that the 5 oligonucleotide GA182 had not annealed to the D19 DNA in the anticipated position at the 3' end of the signal sequence. The DNA cloned in plasmid pMUW1450 contained the first oligonucleotide GA187 in the correct position 5' to the D19 start codon, but the DNA sequence continued 10 past the end of the signal peptide as far as the *Pvu*II site near the center of the gene. Investigation of the reason for the failure of the oligonucleotide GA182 to prime the PCR reaction correctly established that this sequence forms a hair pin loop, so it was unlikely to be 15 available for binding to the D19 gene.

An alternative approach to modifying the 3' end the DNA coding for the PSA signal peptide is described below. Fusion of the promoter with the D19 (PSA) gene, plasmid pMUW1545

20 Plasmid pMUW1450 contains the restriction sites derived from oligonucleotide GA187 that are required for the promoter and D19 gene sequences to be fused. This required a three way ligation to force clone the two DNA fragments into the *Nde*I and *Hind*III sites of pGEM3Z.

25 The DNA fragments to be fused were prepared by cutting 5ug of plasmid pMUW1450 with the restriction enzymes *Nde*I and *Sca*I and then purifying the largest (1.8 Kb) DNA fragment containing the D19 sequences by gel purification as described previously. The *Nde*I cleavage 30 site at the end of this fragment occurs within the D19 sequence coding for the signal peptide. The promoter sequence was prepared by cutting 5ug of plasmid pMUW1480 with the *Hind*III and *Eco*RI restriction enzymes, which cut the *Hind*III site in oligonucleotide GA190 derived sequence 35 5' to the promoter and the *Eco*RI site in the polylinker,

yielding a 0.3 Kb fragment which was then purified by gel electrophoresis. The DNA fragments containing the D19 and promoter fragments were mixed together and digested with the FokI restriction enzyme which creates compatible ends 5 at the ATG start codons in both sequences. The FokI digested fragments were extracted with 50% phenol in chloroform, then chloroform and then precipitated with three volumes of ethanol at -70 degrees. The FokI fragments were ligated with 0.5ug of pGEM3Z which has been 10 cut with HindIII and NdeI, treated with alkaline phosphatase and purified by gel electrophoresis as described for plasmid pMUW1410. The religated plasmids were transformed into the E. coli "Sure" strain as described above and plated onto LB agar containing 100 ug 15 ampicillin per ml. E. coli clones resistant to ampicillin were selected, their plasmids (e.g. pMUW1545) prepared by alkaline lysis and checked for the presence of the BgIII restriction site from the oligonucleotide GA190, and the absence of the NsII restriction sites that should have 20 been removed from both of the inserted fragments.

Construction of the Full Promoter/Signal Sequence, pMUW1594

In order to replace the 3' end of the D19 sequence coding for the signal peptide a synthetic DNA sequence was cloned into the NdeI restriction site of plasmid 25 pMUW1545. The synthetic DNA sequence is composed of two synthetic oligonucleotides GA297 (+ve strand) and a complementary sequence GA296 (-ve stand) which anneal to form double stranded DNA with ends compatible with the NdeI restriction site. In designing these 30 oligonucleotides, the opportunity was taken to change the DNA sequence to optimize the codon usage for highly expressed genes, remove the potential to form hair pin loops and to remove the NdeI restriction site used to insert the oligonucleotides, leaving a single NdeI site 35 suitable for cloning at the signal peptide cleavage site.

- 45 -

The DNA sequence changes do not alter the encoded amino acid sequence of the signal peptide.

The oligonucleotide GA297 and GA296 were phosphorylated with T4 kinase. 50p moles of each 5 oligonucleotides in 50ul "One-for-all" buffer (Pharmacia) and 2uM dATP was incubated with 20 units T4 ligase at 37 degrees for 30 minutes and then the enzyme destroyed by heating to 100 degrees.

Plasmid pMUW1545 was linearized with NdeI restriction 10 enzyme and ligated with 1 p mole of phosphorylated oligonucleotides GA297 and GA296 using 0.5 units of T4 ligase at 4 degrees overnight. The religated plasmids were transformed into the E. coli strain "Sure" (Statagene) and after one hour incubation at 37 degrees 15 the organisms were inoculated into 500 ml LB broth containing 100 ug ampicillin per ml. After being shaken for 18 hours at 37 degrees, the cells were harvested and the mixed population of plasmids purified by alkaline lysis. 5 ug of the resulting mixture of plasmids were 20 digested with the Hind III restriction enzyme and an approximately 0.3 Kb fragment of DNA purified by gel electrophoresis as described above. This 0.3 Kb fragment of DNA could only come from plasmids that are cut in the polylinker and also have the synthetic DNA sequence (which 25 contains a second HindIII site) inserted into the NdeI restriction site of pMUW1545. Thus, this 0.3Kb fragment contains the full promoter - signal sequence construct.

The 0.3 Kb promoter - signal sequence was ligated into pGEM3Z that had been cut with HindIII, treated with 30 alkaline phosphatase and purified by gel electrophoresis. The religated plasmid were transformed into the E. coli strain "Sure" (Statagene) and plated onto LB agar containing 100 ug ampicillin per ml. E. coli clones resistant to ampicillin were selected and their plasmids 35 (e.g. pMUW1594) prepared by alkaline lysis. Plasmids were

checked for size and the correct orientation of the promoter (i.e. 5' to the polylinker) using the position of the BgIII site 5' to the promoter. Clones were further screened by T7 polymerase sequencing (Pharmacia) using 5 oligonucleotide GA187 to check the orientation of the inserted synthetic DNA sequence. Plasmid pMUW1594 had the required promoter and signal sequence in frame with the pGEM3Z polylinker encoded lac operon sequences.

Cloning the Actin 15 Polyadenylation Signal, Plasmid  
10 pMUW1512

The two synthetic oligonucleotides GA189 and GA186 were used as primers to amplify the actin 15 polyadenylation sequence in a polymerase chain reaction (PCR). The two oligonucleotide primers were each designed 15 as two sections, the 5' end of the sequences containing restrictions sites required for cloning and the 3' end of the sequences specially matching the sequence of the Actin 15 gene in plasmid pTS1 (Chang et al (1989) Nucleic Acids Res. 17, 3655-3661). The 3' end of the 20 oligonucleotide GA189 is designed to bind at the stop codon of the actin 15 gene and has one extra base pair added to the original sequence in order to place stop codons in all three reading frames, while the 3' end of oligonucleotide GA186 is complementary to the sequence 25 approximately 305 bp 3', immediately preceding a EcoRV restriction site. The oligonucleotide GA186 replaces the EcoRV restriction site with BgIII and EcoRI site for use in cloning.

The PCR amplification of the polyadenylation sequence 30 was carried out using the identical DNA preparations and methods to the cloning of the actin 15 promoter described above, apart from the use of a different pair of oligonucleotides and the transformation of the plasmids into the "Sure" strain of E. coli. The plasmids produced 35 (e.g. pMUW1512) were digested by the restriction enzyme

• PvuII and screened for the presence of a fragment of approximately 800 bp, comprised of 379 bp of pGEM3Z sequences containing an approximately 400 bp insert. The plasmids were further digested with the restriction  
5 enzymes BglIII, EcoRI and KpnI (separately) to check for the presence of the restriction sites from the two oligonucleotides. Plasmids pMUW1512 and pMUW1515 (opposite orientations of the insert) were sequenced to confirm the polyadenylation signal contained no errors  
10 using a T7 polymerase sequencing kit (Pharmacia) and commercially supplied oligonucleotides (Promega) which anneal to SP6 and T7 regions flanking the polylinker.

5 ug of plasmid pMUW1512 was digested with KpnI and subsequently with EcoRI restriction enzymes and an  
15 approximately 0.4 Kb fragment containing the polyadenylation signal purified by gel electrophoresis as described previously. This 0.4Kb fragment was ligated into 1 ug of plasmid pGEM3Z which was also digested with KpnI and EcoRI, treated with alkaline phosphatase and then  
20 purified by gel electrophoresis. The plasmids were transformed into E. coli strain "Sure" plated onto LB agar containing ampicillin as described previously. Plasmids (e.g. pMUW1560) from the ampicillin resistant clones were screened for the correct sized insert (0.4 Kb) and the  
25 presence of a BglIII site derived from oligonucleotide GA186. Plasmid pMUW1560 contains the actin 15 polyadenylation signal in the correct position and orientation for the final expression cassette.

Construction of the Complete Expression Cassette, Plasmid  
30 pMUW1621

The expression cassette was completed in a single cloning step combining the fused promoter/signal sequence from plasmid pMUW1594 with the polyadenylation sequence from plasmid pMUW1560.

35 Plasmid pMUW1560 was digested with the restriction

enzymes SalI and ScaI and the smaller 1.2 kG Kb fragment containing the polyadenylation signal purified by gel electrophoresis as previously described. Plasmid pMUW1594 was also digested with SalI and ScaI enzymes and the 5 larger 2 Kb fragment containing the promoter and signal sequence purified by gel electrophoresis. The two fragments were pooled, ligated and transformed into the "Sure" strain of E. coli. The identity of the isolated plasmids (e.g. pMUW1621) was confirmed by cutting with the 10 restriction enzyme BglII to produce a 0.7 Kb fragment. This fragment can only be produced by plasmids containing two BglII sites, one derived for the oligonucleotide GA190 used to clone the promoter and the second derived from oligonucleotide GA186 used to clone the polyadenylation 15 signal.

Insertion of the Expression Cassette into the Shuttle Vector

The shuttle vector pMUW1580 was linearized using restriction enzyme BamHI, treated with alkaline 20 phosphatase and purified by gel electrophoresis as previously described. The expression cassette in the 0.7 Kb BglII fragment from plasmid pMUW1621 was also purified by gel electrophoresis and ligated into the linearized plasmid pMUW1580. The ends of the DNA fragments produced 25 by the BglII and BamHI enzymes are compatible, so both restriction sites are destroyed in the ligation. The resulting plasmids produced in the E. coli "Sure" strain were digested with C1aI and HindIII enzymes to screen for the presence of the polylinker in the expression cassette 30 and the orientation of the expression cassette in the plasmid. Plasmids pMUW1630 and 1633 had the opposite orientations of the expression cassette.

Insertion of the GUS Gene into the Expression Vector

The GUS gene is the E. coli gene for the enzyme 35 B-glucuronidase which has been modified by the insertion

of SalI and NcoI restriction enzyme sites at the start codon of the gene, an EcoRI site at the 3' end of the gene and a BamHI site removed from the center of the gene (Jefferson et al (1986) PNAS 83, 8447). Plasmid pRAJ275 containing this construct was purchased from Clontech Laboratories Inc., California, USA.

In order that the GUS gene could be easily sequenced, it was inserted into pGEM3Z. The GUS gene was cut out of plasmid pRAJ175 with the restriction enzymes SalI and 10 EcoRI, purified by gel electrophoresis and ligated into plasmid pGEM3Z which had been cut with the same enzymes, treated with alkaline phosphatase and gel purified. The plasmid with the GUS gene inserted was pMUW1550.

A SmaI restriction site was inserted into the EcoRI site of plasmid pMUW1550 using oligonucleotide GA310 as a linker. Oligonucleotide GA310 was phosphorylated as previously described in the section on the construction of the full promoter/signal sequence. 1 pmole of phosphorylated GA130 was mixed with 0.5 ug of plasmid 20 pMUW1550 which had been cut with the EcoRI restriction enzyme and purified by gel electrophoresis. The mixture was ligated at 4 degrees overnight and then transformed into the E. coli "Sure" strain. The transformants were incubated for one hour in SOC medium and then inoculated 25 into 50 ml of LB broth containing 100ug ampicillin per ml. After shaking at 37 degrees for 18 hours the cells were harvested, plasmids purified and cut with the SmaI restriction enzyme. Only the plasmids containing the oligonucleotide GA130 contain a SmaI site, so the 30 linearized plasmids were purified by gel electrophoresis, religated and transformed back into the E. coli strain "Sure".

1 ug of plasmid pMUW1558 containing a the GUS gene with the SmaI restriction site inserted into the EcoRI 35 site at the 3' end of the gene was cut with the

- 50 -

restriction enzymes SalI and SmaI and the 1.9Kb gene purified by gel electrophoresis. The polylinker in the expression vector pMUW1630 was also cut with the restriction enzymes SalI and SmaI, treated with alkaline 5 phosphatase and purified by gel electrophoresis. The two purified DNA fragments were ligated, transformed into the "Sure" strain of E. coli and plasmids purified from ampicillin resistant clones. Plasmid pMUW1653 contained the GUS gene cloned in frame into the SalI site of the 10 expression vector. This was confirmed by restriction mapping using the sites for NcoI and EcoRI enzymes at the 5' and 3' ends of the GUS Gene respectively. The region of the fusion between the sequence encoding the secretion signal and the 5' end of the GUS gene sequencing plasmid 15 was confirmed by DNA sequencing using a T7 polymerase kit (Pharmacia) and oligonucleotide GA187.

Expression of the GUS gene in *D. discoideum*

The suitability of the expression vector for the expression of recombinant genes was confirmed by 20 transforming 5 ug of the expression plasmid pMUW1653 (containing the E. coli B-glucuronidase gene) and 5ug of plasmid pMUW110 (containing the Ddp2 Rep gene and a G418 resistance marker) into D. discoideum strain AX3K using the calcium phosphate precipitation procedure described 25 previously. After one week under G418 selection, the culture supernatant from the transformants was tested for the presence of the GUS enzyme activity using 1mM p-nitrophenyl-B-D glucuronide substrate in 50mM sodium phosphate pH7.0, 10mM 2- mercaptoethanol and 0.1% Triton 30 X-100. A green colouration indicated the presence of the enzyme B-glucuronidase secreted from D. discoideum. Culture supernatants from cells transformed with the expression vector pMUW1630 did not contain B-glucuronidase.

It will be appreciated by persons skilled in the art 35 that numerous variations and/or modifications may be made

- 51 -

to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as 5 illustrative and not restrictive.

## CLAIMS:

1. A polypeptide which facilitates the extrachromosomal replication of a recombinant plasmid in Dictyostelium spp, the recombinant plasmid including an origin of replication derived from a Ddp2-like plasmid but lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.
2. A polypeptide as claimed in claim 1 in which the recombinant plasmid includes an origin of replication derived from plasmid Ddp2.
3. A polypeptide as claimed in claim 2 in which the polypeptide has an amino acid sequence substantially as shown in Figure 2.
4. A polypeptide as claimed in claim 2 in which the polypeptide is encoded by a DNA sequence substantially as shown in Figure 1 from nucleotide 2378 to nucleotide 5038.
5. A recombinant plasmid vector including an origin of replication derived from plasmid Ddp2 or plasmid pDG1 and lacking functional genes for extrachromosomal replication in wide type Dictyostelium spp.
6. A recombinant plasmid vector is claimed in claim 5 in which the vector includes an origin of replication derived from plasmid Ddp2.
7. A recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1 to nucleotide 2436 or a subsection thereof, and lacking functional genes for extrachromosomal replication in wide type Dictyostelium spp.
8. A recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1153 to nucleotide 1775 or a subsection thereof, and lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.
9. A recombinant plasmid vector containing the DNA sequence TGTCAATGACA but lacking functional genes for

extrachromosomal replication in wild type Dictyostelium spp.

10. A recombinant plasmid vector containing a DNA sequence substantially as shown in Figure 1 from nucleotide 1 to nucleotide 3241 or a portion thereof and lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

11. A recombinant plasmid vector as claimed in any one claims 5 to 10 in which the recombinant plasmid includes a heterologous DNA sequence(s) encoding a desired polypeptide together with a promoter sequence(s) that controls the expression of the heterologous DNA sequence(s).

12. A recombinant plasmid as claimed in claim 11 in which the plasmid includes a DNA sequence encoding a polypeptide signal for secretion of the desired polypeptide.

13. A recombinant plasmid vector as claimed in any one of claims 5 to 10 in which the recombinant plasmid vector includes an expression cassette comprising a promoter DNA sequence derived from Dictyostelium Actin 15 gene, a DNA sequence encoding the secretion signal peptide sequence of the D19 gene of the protein PSA and a DNA signal sequence for RNA polyadenylation derived from the Actin 15 gene.

14. A DNA molecule including a nucleotide sequence which encodes the polypeptide as claimed in any one of claims 1 to 3 and which is capable of transforming Dictyostelium strains such that the recombinant plasmid vectors as claimed in any one claims 5 to 13 are capable of extrachromosomal replication in the transformed Dictyostelium strain.

15. A DNA molecule as claimed in claim 14 in which the DNA molecule includes a sequence substantially as shown in Figure 1 from nucleotide 2378 to nucleotide 5038, or part thereof.

16. A recombinant strain of Dictyostelium spp in which

the recombinant strain includes a gene encoding a polypeptide which facilitates the extrachromosomal replication of a recombinant plasmid, the recombinant plasmid including an origin of replication derived from a Ddp 2-like plasmid but lacking functional genes for extrachromosomal replication in wild type Dictyostelium spp.

17. A recombinant strain of Dictyostelium as claimed in claim 16 in which the recombinant plasmid is as claimed in any one of claims 5 to 13.

18. A recombinant strain of Dictyostelium as claimed in claim 16 or 17 in which the gene encoding the polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid is present in a chromosome of the recombinant strain of Dictyostelium.

19. A recombinant strain of Dictyostelium as claimed in any one of claims 16 to 18 in which the gene encoding the polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid has a DNA sequence substantially as shown in Figure 1 from nucleotide 2378 to nucleotide 5038.

20. A recombinant strain of Dictyostelium as claimed in claim 18 in which the gene encoding the polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid has a DNA sequence substantially as shown in Figure 1 from nucleotide 1885 to nucleotide 5292.

21. A recombinant strain of Dictyostelium as claimed in any one of claims 15 to 20 in which the recombinant strain of Dictyostelium harbours a recombinant plasmid as claimed in any one of claims 5 to 13.

22. A method of producing a desired polypeptide comprising the following steps:-

1. Transforming a recombinant strain of Dictyostelium spp with a recombinant plasmid vector including a DNA sequence encoding the desired

polypeptide and sequences enabling the expression of the DNA sequence encoding the desired polypeptide;

2. Culturing the recombinant strain of Dictyostelium under conditions which allow the expression of the DNA sequence encoding the desired polypeptide and allowing the desired polypeptide to be produced either as a cell bound form or secreted;

3. Recovering the desired polypeptide;

characterised in that the recombinant plasmid vector includes an origin of replication derived from a Ddp2-like plasmid but lacks the functional genes for extrachromosomal replication in wild type Dictyostelium spp, and the recombinant strain of Dictyostelium includes a gene encoding a polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid.

23. A method as claimed in claim 22 in which the desired polypeptide is produced in a cell bound form.

24. A method as claimed in claim 22 in which the gene encoding the polypeptide which facilitates the extrachromosomal replication of the recombinant plasmid is present in the chromosome of the recombinant strain.

25. Recombinant plasmid vector pMUW102 as hereinbefore described.

26. Recombinant plasmid vector pMUW111 as hereinbefore described.

27. Recombinant plasmid vector pMUW110 as hereinbefore described.

28. Recombinant plasmid vector pMUW130 as hereinbefore described.

29. Recombinant plasmid vector pMUW1530 as hereinbefore described.

30. Recombinant plasmid vector pMUW1570 as hereinbefore described.

31. Recombinant plasmid vector pMUW1580 as hereinbefore described.

- 56 -

32. Recombinant plasmid vector pMUW1594 as hereinbefore described.
33. Recombinant plasmid vector pMUW1560 as hereinbefore described.
34. Recombinant plasmid vector pMUW1621 as hereinbefore described.
35. Recombinant plasmid vector pMUW1630 as hereinbefore described.
36. Recombinant plasmid vector pMUW1633 as hereinbefore described.
37. Recombinant plasmid vector pMUW1600 as hereinbefore described.

1/30

Fig. 1-1

SalI

TCGACAAATA TCAAGGGTTG GAATCTTGTA AAAATTTCC CGTTATCGCA  
 10 20 30 40 50

HindIII

AACAATCAAA GTTTAAGCTT CAATCTCAA TAATAATTT AACTTTATCT  
 60 70 80 90 100

ClaI

CTTTCAATTT TAATAATTT TTTCAAAAAT TGAAAATGGT ATAGATCGAT  
 110 120 130 140 150

AGATCACCTT TTTTAGAGAT AAACCATGAA AAAGACATAA AAAATAAAGG  
 160 170 180 190 200

TCATCAAAGT ATTAAAAAAA ATTAATTATC TTTTTAACCTT TGAAAAAAA  
 210 220 230 240 250

AAATAAAAAA AAATAAAAAA AAAAAATTCT TTGTTTAAT AACTTTAAA  
 260 270 280 290 300

ATTATTAAAAA ATAGTATAGA TTTAAAGATC ACAATTTTT ATAATTAACT  
 310 320 330 340 350

ACATAAAATT TATAAAAAAT GAGGGTCATG AAGATATATA AATAATTATT  
 360 370 380 390 400

TAATTATTAA ATATTTAATT ATTTATTTAA CTTAAAAAAA AAAAAAAGGA  
 410 420 430 440 450

AAAAAAGGAA AAAAAAAGTG AAAAAGGTGG GAAAATGAAA AAAAAAAGTGA  
 460 470 480 490 500

AAAAAATGCC CAAAAAAATT TTTATATGAG AAAAAAATTA CGTAAAAAAA  
 510 520 530 540 550

AAATAAGTCT GACCCAAATC GAAAAATAAT AAAAGAGGGG AAAGTAATTA  
 560 570 580 590 600

TAACTAGGTT AGTTTTTAT AATTTTACA TATTGTTAA TAACTTTAA  
 610 620 630 640 650

NdeI

TTTGAAATCA TATGATATTA CATCGTCCCC TTGAAAAAAA AAAAAAAAAT  
 660 670 680 690 700

TTTTTTTCA AACATTTCA TTTTTAAAAA AATGATATAA AATTTAAAC  
 710 720 730 740 750

TAAACTATTT TATTAAATAC AAATATATAA CTTTATCTTA ATCAATTTT  
 760 770 780 790 800

BglII

TTGGTTTATA CATATTTATG TTCTGACTGA AGTATAGATC TTATTACTAA  
 810 820 830 840 850

AGTTCAAAA GTTTAAAAA AAATTAAGG GGGTAAATAT ATAACTTCT  
 860 870 880 890 900

2/30

Fig. 1-2

GTTTTTTCAT ATTCTGTCAT GACAGAAAGG TAAAAAGTGT CATGACAAAA  
 910 920 930 940 950  
 AAAAAAAA AAAAAATTAA TTTCTTCAT AGGTATTGAA ATGACCTCCG  
 960 970 980 990 1000  
 TTTTAATAA AAAGTATATA TTTGTGCTT CCTAGATGAA ATAAGGTTAT  
 1010 1020 1030 1040 1050  
 TTGAGCTTAA TTCAGATTAT TATAAGATTA TTATAAAAAA ATGAAAAACT  
 1060 1070 1080 1090 1100  
 GTCATGACAG TTTTGTAAG TTTCTTATAG TTTTTTTAA TGATCTGAAT  
 1110 1120 1130 1140 1150  
HindIII XbaI  
 TAAGCTTAAA TAACCTTATT TCATCTAGAC GAGCACAAAT ATATACTTT  
 1160 1170 1180 1190 1200  
 TATTAACAC GGAGGTCATT TCAATACCTA TTGAAGAAAT AAATTTTT  
 1210 1220 1230 1240 1250  
 TTTTTTTT TTTGTCACT CACTTTTT TTTTGTCAT GACAGAATTG  
 1260 1270 1280 1290 1300  
 AAAAAACAG AAAGTTATAT ATTTACCCCCC TTTAATTTT TTTAAACTT  
 1310 1320 1330 1340 1350  
BglII  
 TTGAAACTTT AGTAATAAGA TCTATACTTC AGTACGAACA TAAATATGTA  
 1360 1370 1380 1390 1400  
 TAAACCAAAA AAATTGATTA AGATAAAGTT ATATGTTGT ATTTAATAAA  
 1410 1420 1430 1440 1450  
 ATAGTTAGT TAAATTTT ATATCATT TTTAAAAATG AAAATGTTG  
 1460 1470 1480 1490 1500  
NdeI  
 AAAAAAAA TTTTTTTT TTTTTCAAC GGGACGATGT AATATCATAT  
 1510 1520 1530 1540 1550  
 GATTCAAAAT TAAAGTTAT TAACAAATAT GTAAAATTA TAAAAAACTA  
 1560 1570 1580 1590 1600  
 ACCTAGTTAT AATTACTTTC CCCTCTTTT TTTTTTTT TTTGTCACT  
 1610 1620 1630 1640 1650  
 CACTTTTTT TTTTGTCAT GACACTTTT TTTAAAAAAA AAAAAAAA  
 1660 1670 1680 1690 1700  
 ATGTTAAAT ACTATTTGAT GACATTCACT TTTCTAGTT TTTTTTTAGA  
 1710 1720 1730 1740 1750  
ClaI  
 TAGATATAAA AATAAATTGC CTATCGATAT ATACTTAATT TATTAAGATT  
 1760 1770 1780 1790 1800

3/30

Fig. 1-3

GAATAATATT TTAATTTTA ATAAATTCTA CTTTTTTTT TTTTTCTTT  
 1810 1820 1830 1840 1850  
 TTTTTTAAA TTTTAAAATT TTTTTTTTT ATTAGATCTC ATAATTAAAA  
 1860 1870 1880 1890 1900 BglII  
 ATCAATTAA AATTAAAAGT TATTTTAAA TATGCAAAAA CTATAAAAAA  
 1910 1920 1930 1940 1950  
 CTAATGTAGT TTAACCAACT TTTTCTATT TCTTTTTTT TTTTTTTTT  
 1960 1970 1980 1990 2000  
 TTTTACTTT GAAAAAAA AAAAAAAA AAAAAAAA AACCTCATT  
 2010 2020 2030 2040 2050  
 ATAAATATTAA ATTACTTTGG TTTTTTTG A TTTTTTTTT AATAAATTAA  
 2060 2070 2080 2090 2100  
 AAATTTTATT CTCTATCTAA TTATACCTTA TTTATAAATA TTGGAATAAT  
 2110 2120 2130 2140 2150  
 ATATCAAATA TTTATCAGTT TTGGCATGAC AATTTAATT ATATTTATTT  
 2160 2170 2180 2190 2200  
 TTTGATTAGT TTTTTTTTT TTTTTTTTT AAAATTCTT TTTTTTTTT  
 2210 2220 2230 2240 2250  
 TTTATTTTA ATTTTAATT TTTATTTTC CCACACTTC ATTTTATTT  
 2260 2270 2280 2290 2300  
 ATTTTATTTA TTGTAAATTC ATTTTATTTA TTTTAATTAA AATAGTTTG  
 2310 2320 2330 2340 2350  
START  
 GTTTAATTAA ATTCAAAGAT TTTAAAATG GACGAACCTTA TTTCTGGGA  
 2360 2370 2380 2390 2400  
ECORI  
 TAGGTTTTT AAGTTTTTG TAATACTTTT GGAAGAATTC AAAGGTTGTA  
 2410 2420 2430 2440 2450  
 AAAGAAATGA TGTGCGTTTG AGTGTGATT ATGACATTCT TTCTGGTATT  
 2460 2470 2480 2490 2500  
 TATTGCCAC GTACATTTGT ACTAAAGGAA GTCTTAGAG CAGTGGCCGT  
 2510 2520 2530 2540 2550  
 CTCTTATGAT GAATCTGAAA TAGATTTATT CAGATTGGGT TCAGTGTTC  
 2560 2570 2580 2590 2600  
 CTGGTACTTC TTTATATTCA TATATTCCAG GTATTTCAAG TTTAAAAGAT  
 2610 2620 2630 2640 2650  
 TTCCTTTAA TTTCAAAAAC TAAATCGGGT AAAATAAGAG TTTCGGATGT  
 2660 2670 2680 2690 2700

4/30

Fig. 1-4

BclI

AGATCAAGCA	ATATTAATTT	TTGATCATT	TTCTAGAATT	TCAGATAAAC
2710	2720	2730	2740	2750

AAGTATTCG	TAAAGATATT	ATTCCAGGTT	ATAGAACCTT	TGAAAAATCA
2760	2770	2780	2790	2800

PstI

ATATCGAGCG	AGTACAAAAT	CTCGGATGGT	CGTGCTGCAG	GAGTGAGTTG
2810	2820	2830	2840	2850

GTTCAATT	TTAGTAAA	TAAGCACT	TTGTAAAA	CATCCCTTGT
2860	2870	2880	2890	2900

TTGCCGAAA	TCCAACATAT	AAACATGTGG	ATTTTATATC	AATGTTATCA
2910	2920	2930	2940	2950

CTGGTGCATG	GAATCATTGT	TGATTCCCAA	AATGAAGATG	AGAATAATGT
2960	2970	2980	2990	3000

TTCGGCAATG	TACTCTCTGA	ATCCTTTGT	GGATCTTGAA	AAAAGTGATA
3010	3020	3030	3040	3050

TACCAGGGGC	TGTTCAAAGT	AGAGTTACTA	CAAATAGAAC	TAGAGGTTCA
3060	3070	3080	3090	3100

AGGTCTAATT	CCAATTTGAA	TAATCCAACA	ACAACAACAA	CTACTACTAC
3110	3120	3130	3140	3150

CACTACTACA	ACTACCGCAC	CAATTACTAC	TAGAAGTAAA	AGAAAATCTG
3160	3170	3180	3190	3200

ACGACTCTGT	ACAAGAACAA	AGCTCACGAC	AACCAAAAC	<u>XbaI</u> CTCGAGAAAG
3210	3220	3230	3240	3250

TCTGGTTCTC	TTAAGGATGT	CAGAATTAAAC	AATATATCAG	TAGATTCAAG
3260	3270	3280	3290	3300

TTCCAGTGAA	TCTGATGTGA	TTATGTCAGT	TTCAAACCGT	TTAAAATGTT
3310	3320	3330	3340	3350

ATCTTTGGA	AGCAGTTGTA	AACAAAGGAG	AGATCGGTTT	AGAAGTCGTC
3360	3370	3380	3390	3400

AAAGAAGTTT	TAAAAGATTT	ACAGGACAAA	AATTATTCCA	CAGGTTTACT
3410	3420	3430	3440	3450

TGAAAACATT	TTCAATCACA	ACAAGTCTGA	AAGGGTCATA	ACACTTTCAA
3460	3470	3480	3490	3500

GTAGTTTTT	TGAAATTGCT	TCAAAAATTA	ACTATGATGA	AGTTAAGTTC
3510	3520	3530	3540	3550

AGTGAACCTCA	GTATTGATGT	TCTGGAATCG	GCAAAGAGAT	TAACATTGCA
3560	3570	3580	3590	3600

5/30

Fig. 1-5

GAAAAATACA AATATATTAA TTCCAACCAA TAATTTAAA GAAGGTTTG  
 3610 3620 3630 3640 3650

AATTTTATG GGTTCCAATT GTTAATGGTA TTGCTTCAAC TTCTGTCTTT  
 3660 3670 3680 3690 3700

GTTTCACCAA ATAATTATTC AAGTGGTTCA TTTGCAAATG TAGAATCTGC  
 3710 3720 3730 3740 3750

TTTAAAGTTG ATTCACTTT GCATTCTTT AGGAAATATA AATGGTTCC  
 3760 3770 3780 3790 3800  
ClaI

TCTCTATTAG ATCAATTACA TTTGATACAT TTAAATCGAT TACAAAGGAT  
 3810 3820 3830 3840 3850

CTTATTCCAA TGTGAAAAG AATGCTGGAC CTTGAACAAAG GCTTCCGAAA  
 3860 3870 3880 3890 3900

ACTTAGAGAT GCTTGGAAATA ATAGTAATAA AAAATCCAAA GTTCAAGATA  
 3910 3920 3930 3940 3950  
ClaI

GTGATATTAG TGGCATCGAT ACAGAGGATA CAAAGTTGAT ATCATTGTC  
 3960 3970 3980 3990 4000  
ECORV

CACGAGTTA TAAATGATAA TTTATATTTA AAACATATCAA AAGAAGAAGA  
 4010 4020 4030 4040 4050  
AccI

TGGACTAATG CTAGTAGACT TTCCAACATC AACACTTTT ATGAGATACA  
 4060 4070 4080 4090 4100

ATCCAAATAG CATTGATAAC AAAGTTGGTT TCATGTTCCA TTGCCGTTCA  
 4110 4120 4130 4140 4150

GAGATTCAA AGTTCAAAG TTGTAAAAAC CACTCGATAG ATAACCTTGT  
 4160 4170 4180 4190 4200

TTTATCATT ACTCCAAATA ACATTAaaaa TATATCACAG GATAATGAAA  
 4210 4220 4230 4240 4250

ATGAGCTTAA AAAGAAATAT TCGTTGATGG TCAGTGATTT TAGAAATGTT  
 4260 4270 4280 4290 4300

CCAAAGGTGA CACCAAAATT TATACTTCT GAATTTAAA GGTTTACAAT  
 4310 4320 4330 4340 4350

CATTACGTTCAAC ACACAAATT CATAACAATGC CAATAGAGTA TTTGCCTTGC  
 4360 4370 4380 4390 4400

ACGACATCTC AAGTGGAAATT TCAATCACAA ATGTAAAAA TATCCACGCA  
 4410 4420 4430 4440 4450  
HindII

AAGGGTCAAC GAAACTTTGA AATCTACGAA ACATTACTGG GAAGTACCAAG  
 4460 4470 4480 4490 4500

6/30

Fig. 1-6

GATTATTTCGT GCATTTTCT GCGCTCCATG CTTGATCCAA ATCAATAATT  
 4510 4520 4530 4540 4550  
  
 TTAAATTTGC CACAGATAAG TTAATTGATG ACCAAAGTGT AAATCACCGAG  
 4560 4570 4580 4590 4600  
  
 ATTGCATCTT TGGAAATTAA AAACTTATCA TATCTTCCGC TCGACATCAA  
 4610 4620 4630 4640 4650  
  
 GGTTAGAGGT AGTACAGTTG GAACGATTAA GGGTGGAGAG ACAGCTCCTA  
 4660 4670 4680 4690 4700  
  
 TTATTATAAA CTCAGAAGAA TTTACGTTT CTATCTCATG CCTTGATATT  
 4710 4720 4730 4740 4750  
  
 AGATTTAGTG CATCCTTAAT TTCTAAAACA AAACTAAGCC AACTTCCAAC  
 4760 4770 4780 4790 4800  
  
 ATTTGCTCCA GATGAAAGGT ACAATAAAGA GACTAACATT TTAAAAGTTT  
 4810 4820 4830 4840 4850  
  
 TGGATCAATG TGATGAACCTT ACTCGAACGT TTTAAATAA CTATAAAATA  
 4860 4870 4880 4890 4900  
  
 GCTAATAAAC TATCAACCAT TGAAAATTAT TTATATAATA ATTTTATGGG  
 4910 4920 4930 4940 4950  
  
 ACTAGAAGAT GAAGATGAAG ATGAAGATGA AGATGAAGAT GAAGATGAAG  
 4960 4970 4980 4990 5000  
STOP  
 ATGAAGATGA AGATGAAGAT GAAGACGAAG ATGGGTATTG AATTATCATA  
 5010 5020 5030 5040 5050  
  
 CTTTAAAAAT TAATTAAATA AATAAAAAAA AAAAAATGAT TTCAATTAA  
 5060 5070 5080 5090 5100  
  
 ATATATACAT ATATATATAT ATAAAATGAG ATTAATAAAA CTTTGAGAC  
 5110 5120 5130 5140 5150  
  
 CAACATTAA TGAGATTCT GATGCTGTT ATTGCTGGAATGAGAGC  
 5160 5170 5180 5190 5200  
  
 AAAAGGCTAA AAAACATGAG AGAGAAATATA ATAATAAAGG AAAACTTGGG  
 5210 5220 5230 5240 5250  
ScaI  
 AAAAAGGATC TAGTATCCAT TTCCATATTA ATCCGTGCAG TACTATTAAT  
 5260 5270 5280 5290 5300  
  
 TAAAAAAATA CTTTAAAAAA AATTTTAAAA ACATGGAAAA TTATATAGAT  
 5310 5320 5330 5340 5350  
Clai  
 CGATAGATCA CTAATTTTA AAATTAATA TATTAATTT ATAAAATTG  
 5360 5370 5380 5390 5400

7/30

Fig. 1-7

AAGTTCATCA AGATATATAG ATAATTATTT AATTATTTGA ATTTTTAAAA  
5410 5420 5430 5440 5450

AAAAAAAAAA AAAAAGAAAAA AAAATCAAAT ATGTTTATTG TTTTAAGATT  
5460 5470 5480 5490 5500

ClaI  
TTTTAATCTC GTCAATGATT TTAAAATAAA AATCGATACA TAATTTAAA  
5510 5520 5530 5540 5550

AAAAACCCCTT TACATTTTTT ATTTTAATTG CAAATTATAA CATTTTTAT  
5560 5570 5580 5590 5600

TTTTTTTTTT TTTTTTTTTT TTTTTTTTAA TTTAAATTTT TTTTTTTTTT  
5610 5620 5630 5640 5650

TTTTTTTTAT TTATTTAAAAA TTTAATTATT AATTTTATAA ATAAAAAAATA  
5660 5670 5680 5690 5700

GAAATATAAG TAAAAAAACA AACAAACAAAT AACATATATA AAAAAATACA  
5710 5720 5730 5740 5750

AATAACAAAT AATTAAATAA ATAAATAAAC CATAAAAAT GTATATTAAT  
5760 5770 5780 5790 5800

BglII ScaI  
AAATTTAAAAA GATCTTTATT AGTACTATTG TTACTTTGTA ATATTCTTCC  
5810 5820 5830 5840 5850

Sali  
TG

8/30

Fig. 2-1

REP GENE:

2386	2395	2404	2413	2422
ATG GAC GAA CTT ATT TCT TGG GAT AGG TTT TTT AAG TTT TTT GTA				
m d e l i s w d r f f k f f v				
2431	2440	2449	2458	2467
ATA CTT TTG GAA GAA TTC AAA GGT TGT AAA AGA AAT GAT GTG CGT				
i l l e e f k g c k r n d v r				
2476	2485	2494	2503	2512
TTG AGT GTC GAT TAT GAC ATT CTT TCT GGT ATT TAT TCG CCA CGT				
l s v d y d i l s g i y s p r				
2521	2530	2539	2548	2557
ACA TTT GTA CTA AAG GAA GTC TTT AGA GCA GTG GCC GTC TCT TAT				
t f v l k e v f r a v a v s y				
2566	2575	2584	2593	2602
GAT GAA TCT GAA ATA GAT TTA TTC AGA TTG GGT TCA GTG TTT CCT				
d e s e i d l f r l g s v f p				
2611	2620	2629	2638	2647
GGT ACT TCT TTA TAT TCA TAT ATT CCA GGT ATT TTC AGT TTA AAA				
g t s l y s y i p g i f s l k				
2656	2665	2674	2683	2692
GAT TTC CTT TTA ATT TCA AAA ACT AAA TCG GGT AAA ATA AGA GTT				
d f l l i s k t k s g k i r v				
2701	2710	2719	2728	2737
TCG GAT GTA GAT CAA GCA ATA TTA ATT TTT GAT CAT TTT TCT AGA				
s d v d q a i l i f d h f s r				
2746	2755	2764	2773	2782
ATT TCA GAT AAA CAA GTA TTT CGT AAA GAT ATT ATT CCA GGT TAT				
i s d k q v f r k d i i p g y				
2791	2800	2809	2818	2827
AGA ACC TTT GAA AAA TCA ATA TCG AGC GAG TAC AAA ATC TCG GAT				
r t f e k s i s s e y k i s d				
2836	2845	2854	2863	2872
GGT CGT GCT GCA GGA GTG AGT TGG TTC AAT TTA GTT AGT AAA ATA				
g r a a g v s w f n l v s k i				
2881	2890	2899	2908	2917
AGC ACT TAT TGT AAA AAT CAT CCC TTG TTT GCC GAA AAT CCA ACA				
s t y c k n h p l f a e n p t				
2926	2935	2944	2953	2962
TAT AAA CAT GTG GAT TTT ATA TCA ATG TTA TCA CTG GTG CAT GGA				
y k h v d f i s m l s l v h g				

9/30

Fig. 2-2

2971	2980	2989	2998	3007
ATC ATT GTT GAT TCC CAA AAT GAA GAT GAG AAT AAT GTT TCG GCA				
i i v d s q n e d e n n v s a				
3016	3025	3034	3043	3052
ATG TAC TCT CTG AAT CCT TTT GTG GAT CTT GAA AAA AGT GAT ATA				
m y s l n p f v d l e k s d i				
3061	3070	3079	3088	3097
CCA GGG GCT GTT CAA AGT AGA GTT ACT ACA AAT AGA ACT AGA GGT				
p g a v q s r v t t n r t r g				
3106	3115	3124	3133	3142
TCA AGG TCT AAT TCC AAT TTG AAT AAT CCA ACA ACA ACA ACA ACT				
s r s n s n l n n p t t t t t				
3151	3160	3169	3178	3187
ACT ACT ACC ACT ACA ACT ACC GCA CCA ATT ACT ACT AGA AGT				
t t t t t t t a p i t t r s				
3196	3205	3214	3223	3232
AAA AGA AAA TCT GAC GAC TCT GTA CAA GAA CAA AGC TCA CGA CAA				
k r k s d d s v q e q s s r q				
3241	3250	3259	3268	3277
CCA AAA ACC TCG AGA AAG TCT GGT TCT CTT AAG GAT GTC AGA ATT				
p k t s r k s g s l k d v r i				
3286	3295	3304	3313	3322
AAC AAT ATA TCA GTA GAT TCA AGT TCC AGT GAA TCT GAT GTG ATT				
n n i s v d s s s e s d v i				
3331	3340	3349	3358	3367
ATG TCA GTT TCA AAC CGT TTA AAA TGT TAT CTT TTG GAA GCA GTT				
m s v s n r l k c y l l e a v				
3376	3385	3394	3403	3412
GTA AAC AAA GGA GAG ATC GGT TTA GAA GTC GTC AAA GAA GTT TTA				
v n k g e i g l e v v k e v l				
3421	3430	3439	3448	3457
AAA GAT TTA CAG GAC AAA AAT TAT TCC ACA GGT TTA CTT GAA AAC				
k d l q d k n y s t g l l e n				
3466	3475	3484	3493	3502
ATT TTC AAT CAC AAC AAG TCT GAA AGG GTC ATA ACA CTT TCA AGT				
i f n h n k s e r v i t l s s				
3511	3520	3529	3538	3547
AGT TTT TTT GAA ATT GCT TCA AAA ATT AAC TAT GAT GAA GTT AAG				
s f f e i a s k i n y d e v k				
3556	3565	3574	3583	3592
TTC AGT GAA CTC AGT ATT GAT GTT CTG GAA TCG GCA AAG AGA TTA				
f s e l s i d v l e s a k r l				

10/30

Fig. 2-3

3601	3610	3619	3628	3637										
ACA	TTC	GAG	AAA	AAT	ACA	AAT	ATA	TTA	ATT	CCA	ACC	AAT	AAT	TTT
t	f	e	k	n	t	n	i	l	i	p	t	n	n	f
3646	3655	3664	3673	3682										
AAA	GAA	GGT	TTT	GAA	TTT	TTA	TGG	GTT	CCA	ATT	GTT	AAT	GGT	ATT
k	e	g	f	e	f	l	w	v	p	i	v	n	g	i
3691	3700	3709	3718	3727										
GCT	TCA	ACT	TCT	GTC	TTT	GTT	TCA	CCA	AAT	AAT	TAT	TCA	AGT	GGT
a	s	t	s	v	f	v	s	p	n	n	y	s	s	g
3736	3745	3754	3763	3772										
TCA	TTT	GCA	AAT	GTA	GAA	TCT	GCT	TTA	AAG	TTG	ATT	CAT	CTT	TGC
s	f	a	n	v	e	s	a	l	k	l	i	h	l	c
3781	3790	3799	3808	3817										
ATT	TCT	TTA	GGA	AAT	ATA	AAT	GGT	TTC	CTC	TCT	ATT	AGA	TCA	ATT
i	s	l	g	n	i	n	g	f	l	s	i	r	s	i
3826	3835	3844	3853	3862										
ACA	TTT	GAT	ACA	TTT	AAA	TCG	ATT	ACA	AAG	GAT	CTT	ATT	CCA	ATG
t	f	d	t	f	k	s	i	t	k	d	l	i	p	m
3871	3880	3889	3898	3907										
TCG	AAA	AGA	ATG	CTG	GAC	CTT	GAA	CAA	GGC	TTC	CGA	AAA	CTT	AGA
s	k	r	m	l	d	l	e	q	g	f	r	k	l	r
3916	3925	3934	3943	3952										
GAT	GCT	TGG	AAT	AAT	AGT	AAT	AAA	AAA	TCC	AAA	GTT	CAA	GAT	AGT
d	a	w	n	n	s	n	k	k	s	k	v	q	d	s
3961	3970	3979	3988	3997										
GAT	ATT	AGT	GGC	ATC	GAT	ACA	GAG	GAT	ACA	AAG	TTG	ATA	TCA	TTT
d	i	s	g	i	d	t	e	d	t	k	l	i	s	f
4006	4015	4024	4033	4042										
GTC	CAC	GAG	TTT	ATA	AAT	GAT	AAT	TTA	TAT	TTA	AAA	CTA	TCA	AAA
v	h	e	f	i	n	d	n	l	y	1	k	l	s	k
4051	4060	4069	4078	4087										
GAA	GAA	GAT	GGA	CTA	ATG	CTA	GTA	GAC	TTT	CCA	ACA	TCA	ACA	CTT
e	e	d	g	l	m	l	v	d	f	p	t	s	t	1
4096	4105	4114	4123	4132										
TTT	ATG	AGA	TAC	AAT	CCA	AAT	AGC	ATT	GAT	AAC	AAA	GTT	GGT	TTC
f	m	r	y	n	p	n	s	i	d	n	k	v	g	f
4141	4150	4159	4168	4177										
ATG	TTC	CAT	TGC	CGT	TCA	GAG	ATT	TCA	AAG	TTT	CAA	AGT	TGT	AAA
m	f	h	c	r	s	e	i	s	k	f	q	s	c	k
4186	4195	4204	4213	4222										
AAC	CAC	TCG	ATA	GAT	AAC	CTT	GTT	TTA	TCA	TTT	ACT	CCA	AAT	AAC
n	h	s	i	d	n	l	v	1	s	f	t	p	n	n

11/30

Fig. 2-4

4231	4240	4249	4258	4267
ATT AAA AAT ATA TCA CAG GAT AAT GAA AAT GAG CTT AAA AAG AAA				
i k n i s s q d n e n e l k k k				
4276	4285	4294	4303	4312
TAT TCG TTG ATG GTC AGT GAT TTT AGA AAT GTT CCA AAG GTG ACA				
y s l m v s d f r n v p k v t				
4321	4330	4339	4348	4357
CCA AAA TTT ATA CCT TCT GAA TTT AAA AGG TTT ACA ATC ATT ACG				
p k f i p s e f k r f t i i t				
4366	4375	4384	4393	4402
TTC ACA AAC AAT TCA TAC AAT GCC AAT AGA GTA TTT GCG TTT GAC				
f t n n s y n a n r v f a f d				
4411	4420	4429	4438	4447
GAC ATC TCA AGT GGA ATT TCA ATC ACA AAT GTT AAA AAT ATC CAC				
d i s s g i s i t n v k n i h				
4456	4465	4474	4483	4492
GCA AAG GGT CAA CGA AAC TTT GAA ATC TAC GAA ACA TTA CTG GGA				
a k g q r n f e i y e t l l g				
4501	4510	4519	4528	4537
AGT ACC AGG ATT ATT CGT GCA TTT TTC TGC GCT CCA TGC TTG ATC				
s t r i i r a f f c a p c l i				
4546	4555	4564	4573	4582
CAA ATC AAT AAT TTT AAA TTT GCC ACA GAT AAG TTA ATT GAT GAC				
q i n n f k f a t d k l i d d				
4591	4600	4609	4618	4627
CAA AGT GTA AAT CAC CAG ATT GCA TCT TTG GAA ATT AAA AAC TTA				
q s v n h q i a s l e i k n l				
4636	4645	4654	4663	4672
TCA TAT CTT CCG CTC GAC ATC AAG GTT AGA GGT AGT ACA GTT GGA				
s y l p l d i k v r g s t v g				
4681	4690	4699	4708	4717
ACG ATT AAG GGT GGA GAG ACA GCT CCT ATT ATT ATA AAC TCA GAA				
t i k g g e t a p i i n s e				
4726	4735	4744	4753	4762
GAA TTT ACG TTT TCT ATC TCA TGC CTT GAT ATT AGA TTT AGT GCA				
e f t f s i s c l d i r f s a				
4771	4780	4789	4798	4807
TCC TTA ATT TCT AAA ACA AAA CTA AGC CAA CTT CCA ACA TTT GCT				
s l i s k t k l s q l p t f a				
4816	4825	4834	4843	4852
CCA GAT GAA AGG TAC AAT AAA GAG ACT AAC ATT TTA AAA GTT TTG				
p d e r y n k e t n i l k v l				

12/30

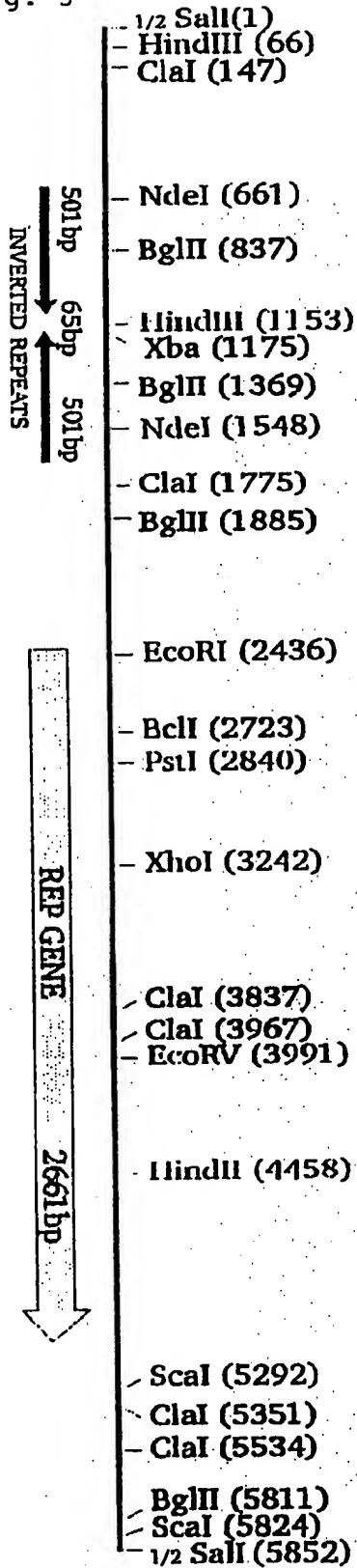
Fig. 2-5

4861	4870	4879	4888	4897
GAT CAA TGT GAT GAA CTT ACT CGA ACG TTT TTA AAT AAC TAT AAA				
d q c d e l t r t f l n n y k				
4906	4915	4924	4933	4942
ATA GCT AAT AAA CTA TCA ACC ATT GAA AAT TAT TTA TAT AAT AAT				
i a n k l s t i e n y l y n n				
4951	4960	4969	4978	4987
TTT ATG GGA CTA GAA GAT GAA GAT GAA GAT GAA GAT GAA GAT GAA				
f m g l e d e d e d e d e d e d e				
4996	5005	5014	5023	5032
GAT GAA GAT GAA GAT GAA GAT GAA GAT GAA GAT GAA GAC GAA GAT				
d e d e d e d e d e d e d e d e d				

GGG TAT  
g y

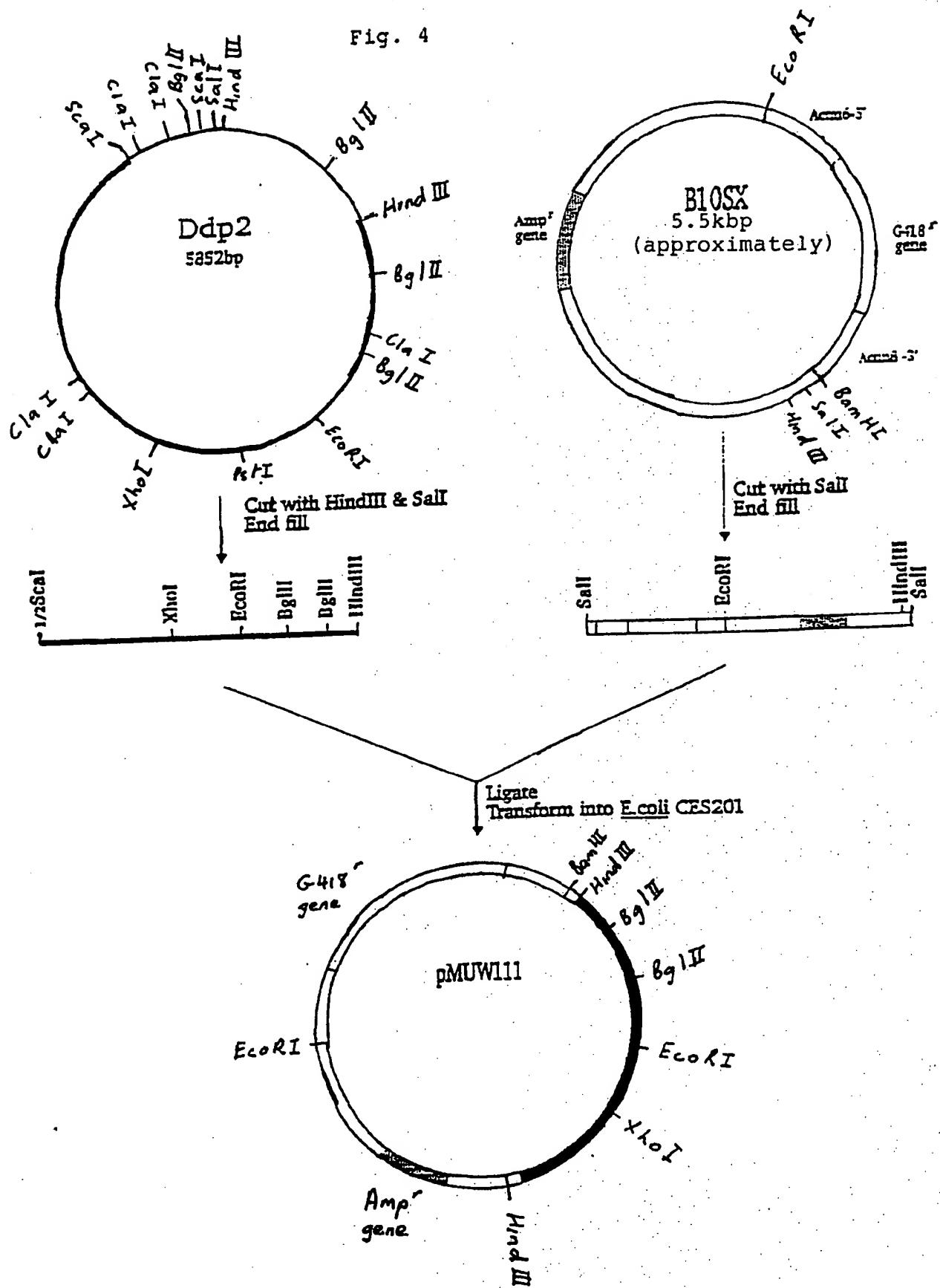
13/30

Fig. 3



14/30

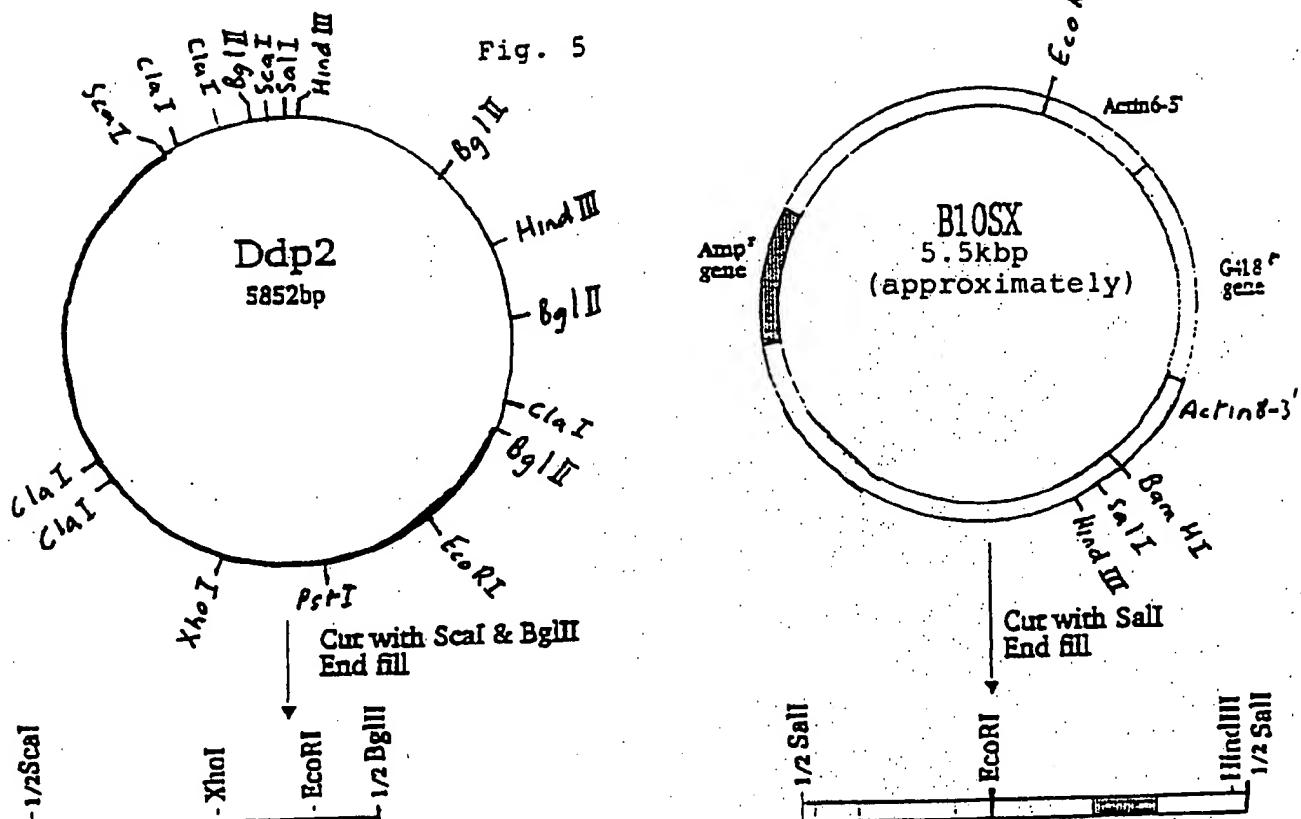
Fig. 4



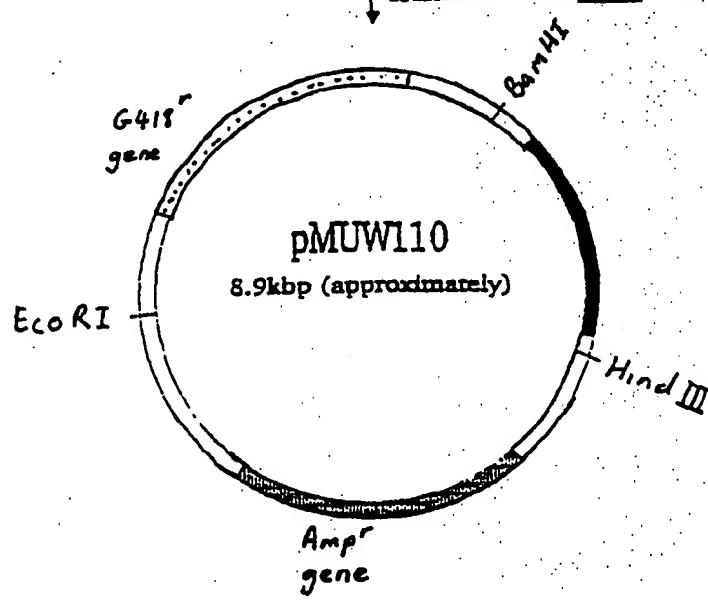
SUBSTITUTE SHEET

15/30

Fig. 5



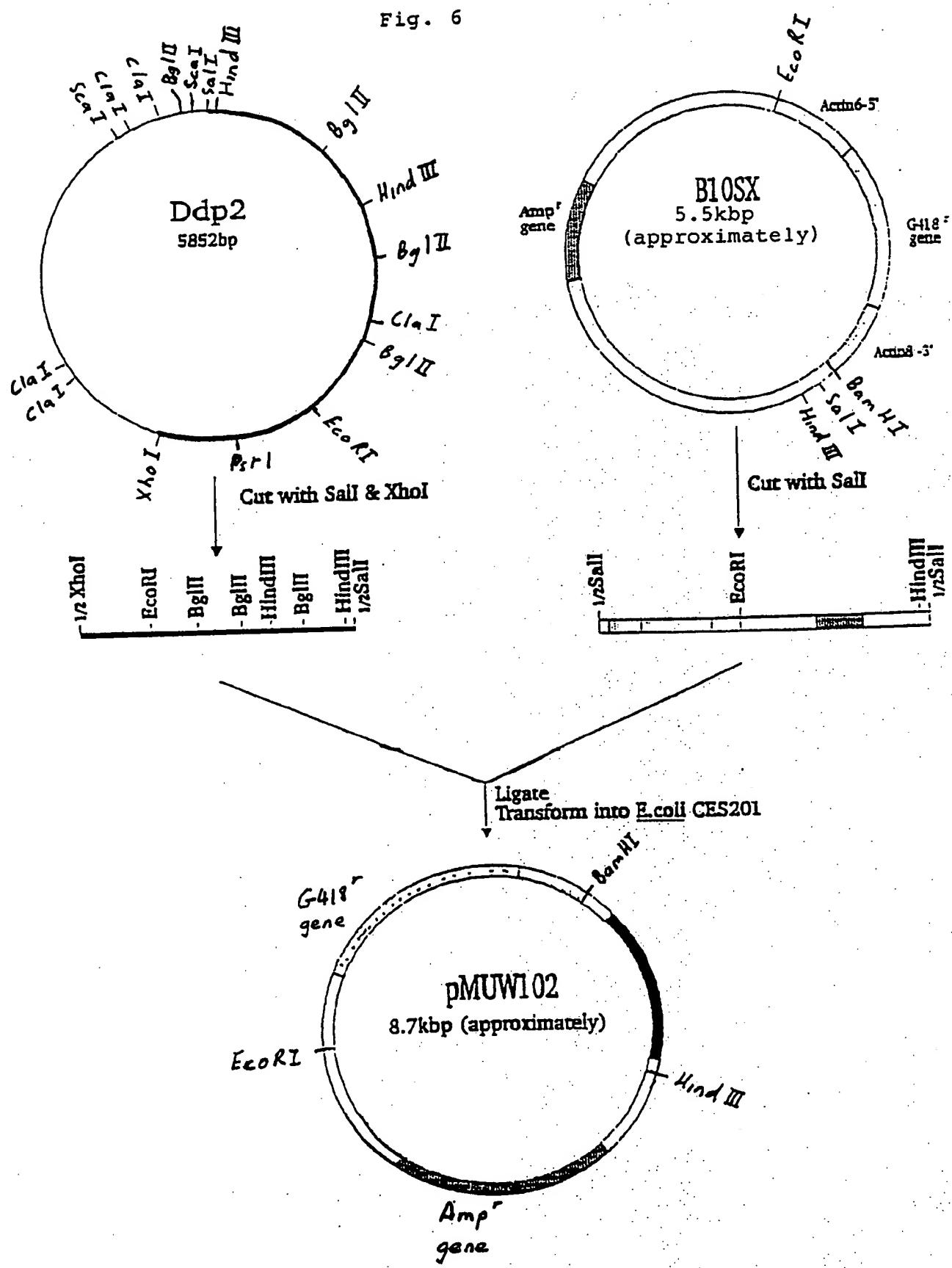
Ligate  
Transform into E.coli CES201



## **SUBSTITUTE SHEET**

16/30

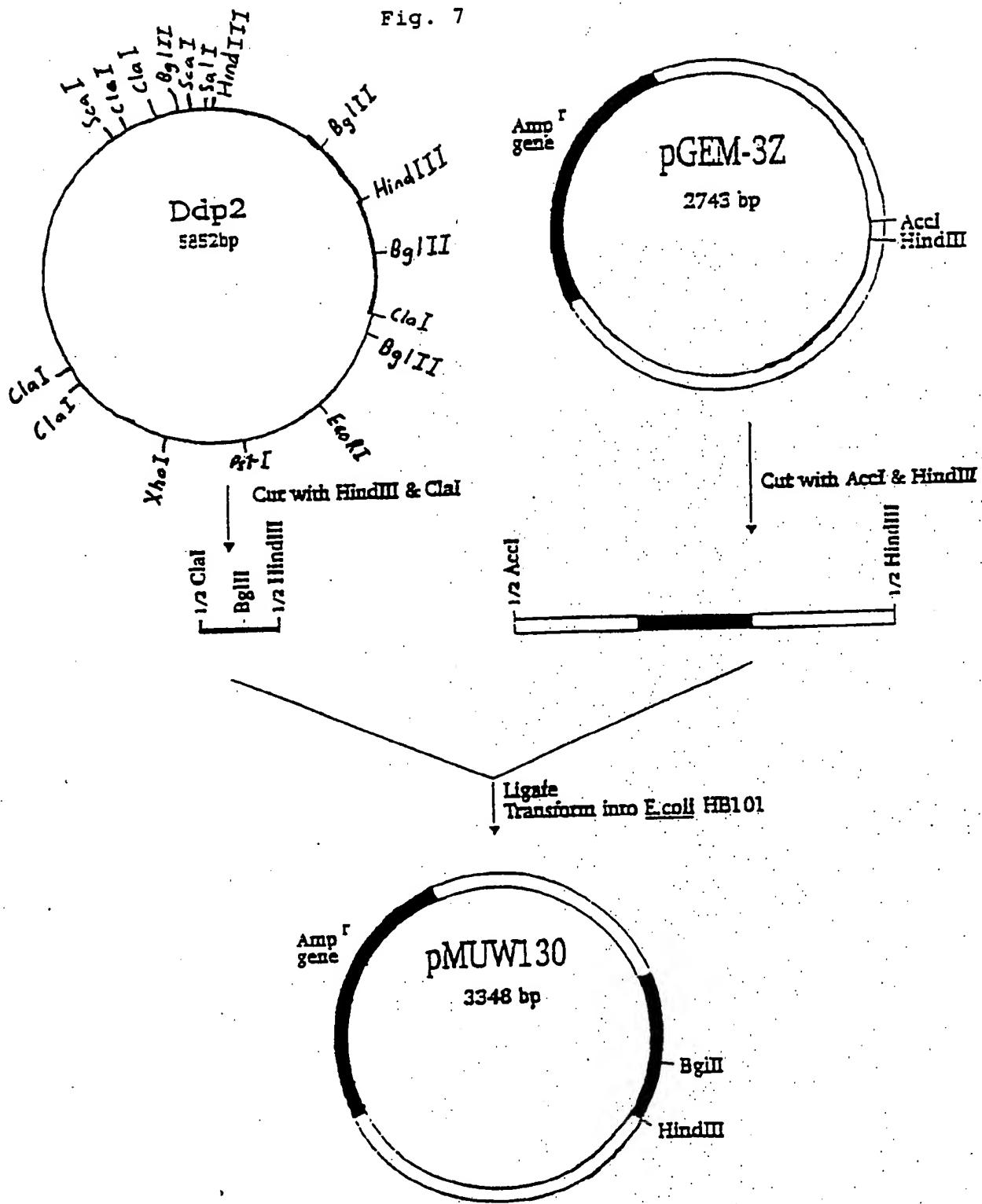
Fig. 6



SUBSTITUTE SHEET

17/30

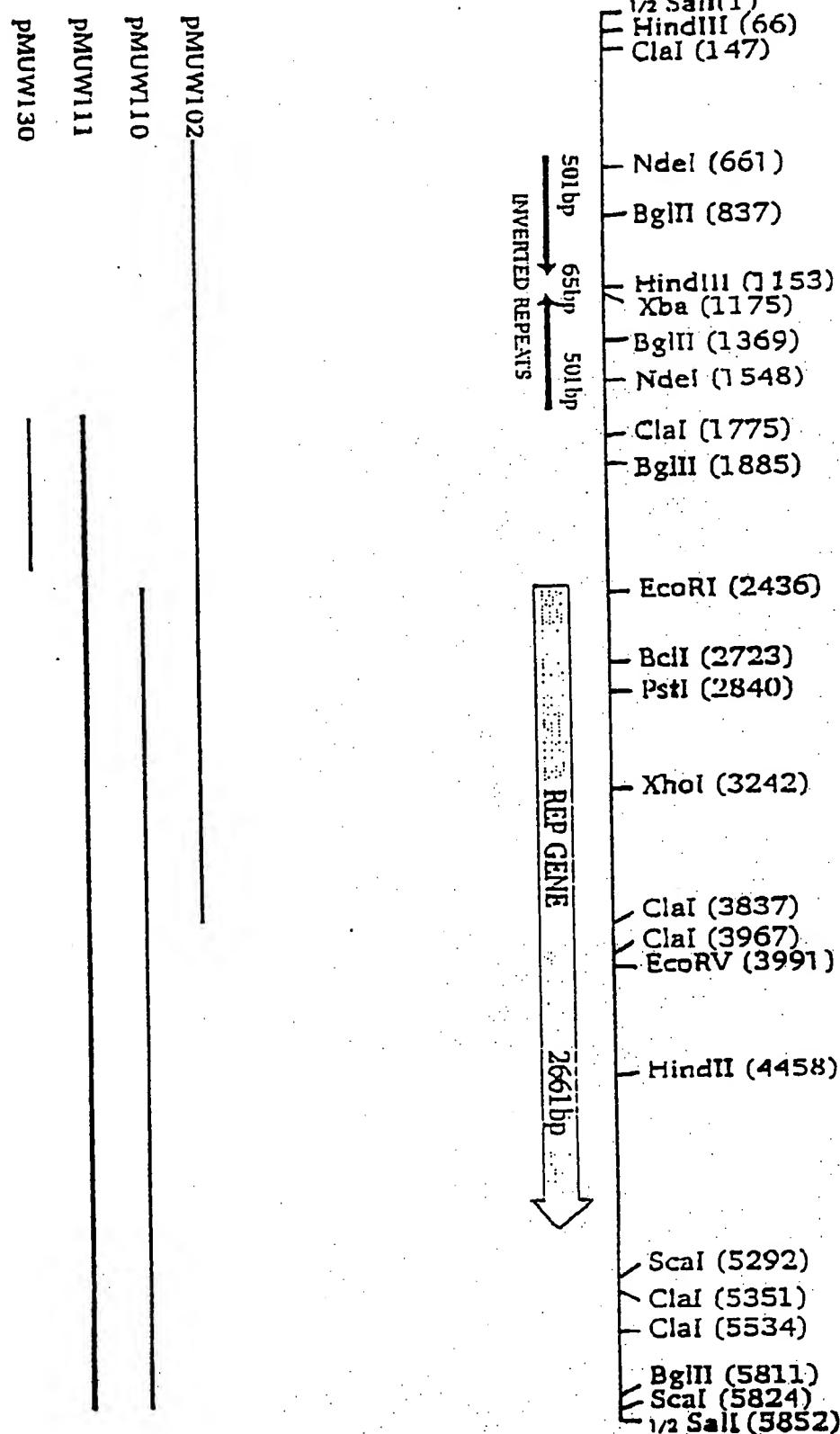
Fig. 7



**SUBSTITUTE SHEET**

18/30

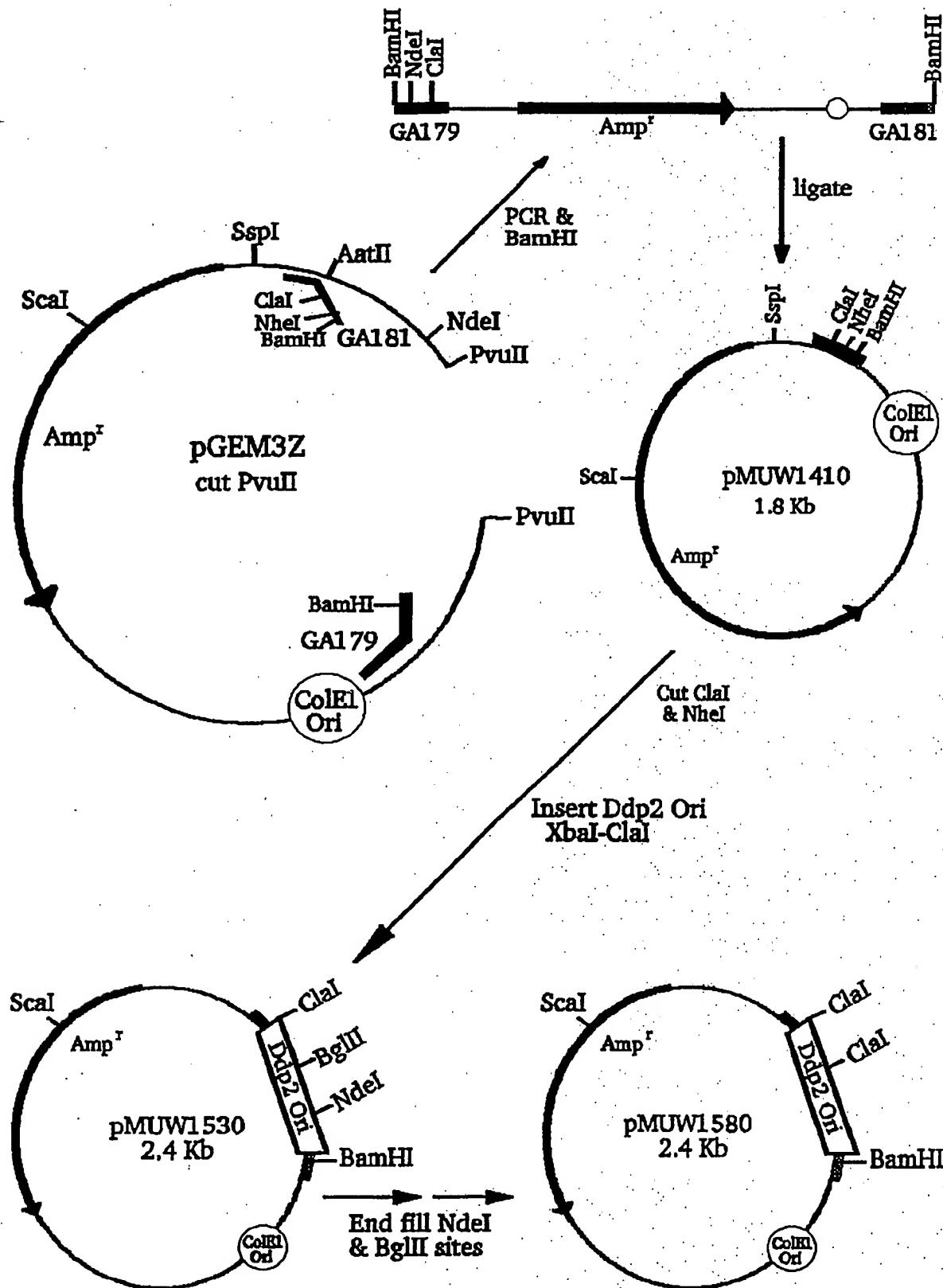
Fig. 8



SUBSTITUTE SHEET

19/30

Fig. 9



SUBSTITUTE SHEET

20/30

Fig. 10-1

Clal	10	20	30	40	50
CGATAGGTGG	CACTTTCCG	GGAAATGTG	GCGGAACCCC	TATTTGTTA	
60	70	80	90	100	
TTTTCTAAA	TACATTCAAA	TATGTATCCG	CTCATGAGAC	AATAACCCTG	
110	120	130	140	150	
ATAAATGCTT	CAATAATATT	GAAAAAGGAA	GAGTATGAGT	ATTCAACATT	
160	170	180	190	200	
TCCGTGTCGC	CCTTATTCCC	TTTTTGC GG	CATTTGCCT	TCCTGTTTT	
210	220	230	240	250	
GCTCACCCAG	AAACGCTGGT	GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG	
260	270	280	290	300	
TGCACGAGTG	GGTTACATCG	AACTGGATCT	CAACAGCGGT	AAGATCCTTG	
310	320	330	340	350	
AGAGTTTCG	CCCCGAAGAA	CGTTTCCAA	TGATGAGCAC	TTTTAAAGTT	
360	370	380	390	400	
CTGCTATGTG	GCGCGGTATT	ATCCCGTATT	GACGCCGGC	AAGAGCAACT	
410	420	430	440	ScaI	450
CGGTCGCCGC	ATACACTATT	CTCAGAATGA	CTTGGTTGAG	TACTCACCAG	
460	470	480	490	500	
TCACAGAAAA	GCATCTTACG	GATGGCATGA	CAGTAAGAGA	ATTATGCAGT	
510	520	530	540	550	
GCTGCCATAA	CCATGAGTGA	TAACACTGCG	GCCAACCTAC	TTCTGACAAC	
560	570	580	590	600	
GATCGGAGGA	CCGAAGGAGC	TAACCGCTTT	TTTGCACAAAC	ATGGGGGATC	
610	620	630	640	650	
ATGTAACTCG	CCTTGATCGT	TGGGAACCGG	AGCTGAATGA	AGCCATACCA	
660	670	680	690	700	
AACGACGAGC	GTGACACCAC	GATGCCGTGA	GCAATGCCAA	CAACGTTGCG	
710	720	730	740	750	
CAAACATTAA	ACTGGCGAAC	TACTTACTCT	AGCTTCCCGG	CAACAATTAA	
760	770	780	790	800	
TAGACTGGAT	GGAGGGGGAT	AAAGTTGCAG	GACCACTTCT	GCGCTCGGCC	
810	820	830	840	850	
CTTCCGGCTG	GCTGGTTTAT	TGCTGATAAA	TCTGGAGCCG	GTGAGCGTGG	
860	870	880	890	900	
GTCTCGCGGT	ATCATTGCAG	CACTGGGGCC	AGATGGTAAG	CCCTCCCGTA	

21/30

Fig. 10-2

910	920	930	940	950
TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT				
960	970	980	990	1000
AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC				
1010	1020	1030	1040	1050
AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA CTTCATTTTT				
1060	1070	1080	1090	1100
AATTTAAAAG GATCTAGGTG AAGATCCTT TTGATAATCT CATGACCAAA				
1110	1120	1130	1140	1150
ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA				
1160	1170	1180	1190	1200
GATCAAAGGA TCTTCTTGAG ATCCTTTTT TCTGCGCGTA ATCTGCTGCT				
1210	1220	1230	1240	1250
TGCACAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTGTT GCCGGATCAA				
1260	1270	1280	1290	1300
GAGCTACCAA CTCTTTTCC GAAGGTAAC GGCTTCAGCA GAGCGCAGAT				
1310	1320	1330	1340	1350
ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA				
1360	1370	1380	1390	1400
ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAACCT GTTACCAGTG				
1410	1420	1430	1440	1450
GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG				
1460	1470	1480	1490	1500
ATAGTTACCG GATAAGGCAGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA				
1510	1520	1530	1540	1550
CACAGCCCCAG CTTGGAGCGA ACGACCTACÀ CCGAACTGAG ATACCTACAG				
1560	1570	1580	1590	1600
CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGCGGACAG				
1610	1620	1630	1640	1650
GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC				
1660	1670	1680	1690	1700
CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC				
1710	1720	1730	1740	1750
TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATC				
1760	1770	1780	1790	1800
GAAAAACGCC AGCAACGCAGG CCTTTTACG GTTCCTGGCC TTTTGCTGGC				

22/30

Fig. 10-3

1810 BamHI 1830 1840 1850  
CTTTGCTGG CCTTGGATC CGCTAGACGA GCACAAATAT ATACTTTTA

1860 1870 1880 1890 1900  
TTAAAAACGG AGGTCAATTTC AATACCTATT GAAGAAATAA ATTTTTTTT

1910 1920 1930 1940 1950  
TTTTTTTTT TGTCATGACA CTTTTTTTT TTTGTCAATGA CAGAATTGAA

1960 1970 1980 1990 2000  
AAAAACAGAA AGTTATATAT TTACCCCCCTT TAATTTTTT TAAACTTTT

2010 BglII 2030 2040 2050  
GAAACTTTAG TAATAAGATC TATACTTCAG TACGAACATA AATATGTATA

2060 2070 2080 2090 2100  
AACCAAAAAA ATTGATTAAG ATAAAGTTAT ATGTTGTAT TTAATAAAAT

2110 2120 2130 2140 2150  
AGTTTAGTTT AAAATTTTAT ATCATTTTT AAAAAATGAA AATGTTGAA

2160 2170 2180 2190 NdeI 2200  
AAAAAAAATT TTTTTTTTTT TTTTCAACGG GACGATGTAA TATCATATGA

2210 2220 2230 2240 2250  
TTCAAAATTA AAAGTTATTA ACAAAATATGT AAAAATTATA AAAAACTAAC

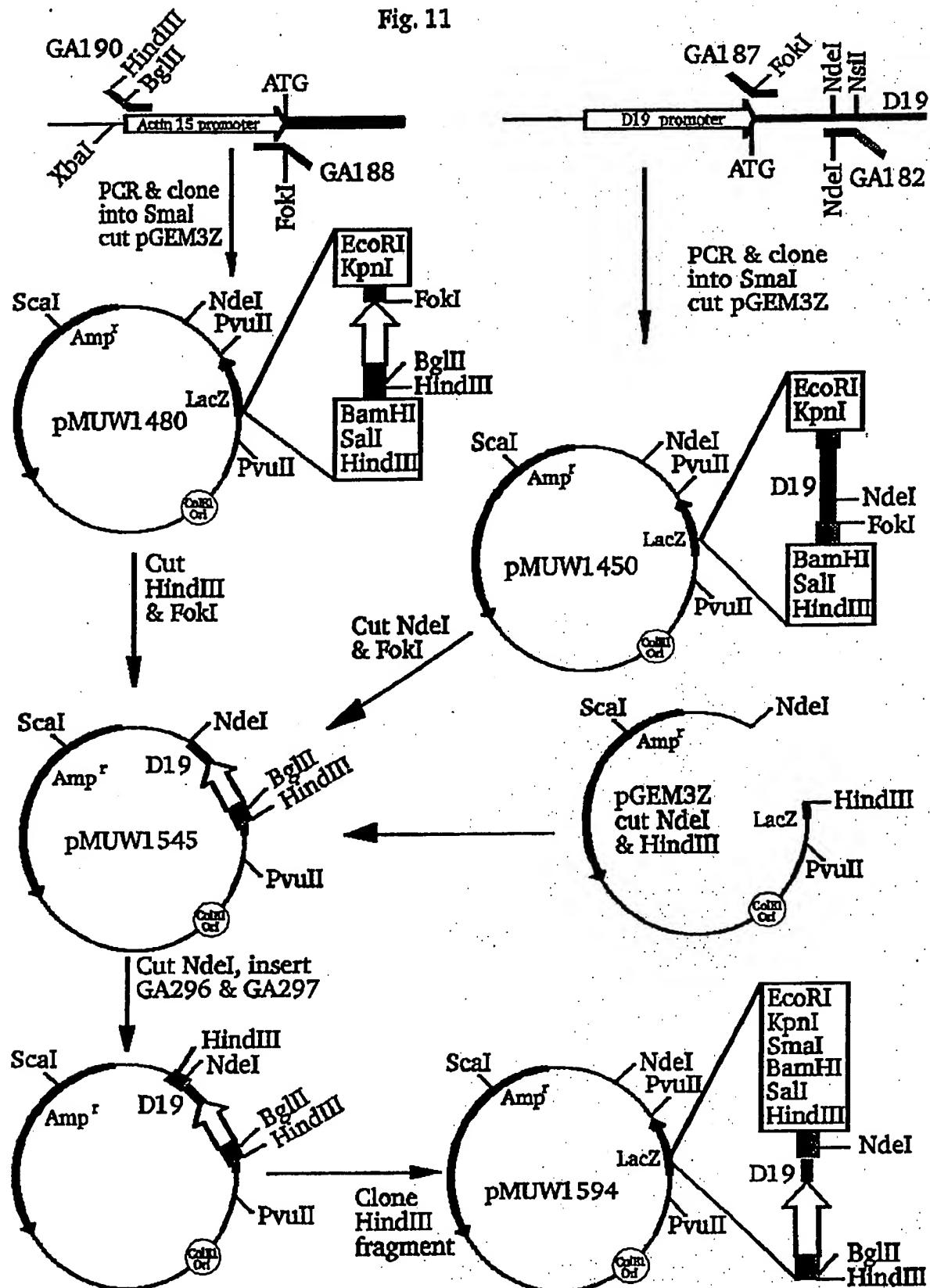
2260 2270 2280 2290 2300  
CTAGTTATAA TTACTTTCCC CTCTTTTTT TTTTTTTTT TGTCAATGACA

2310 2320 2330 2340 2350  
CTTTTTTTTT TTTGTCAATGA CACTTTTTT TTAAAAAAA AAAAAAAAAT

2360 2370 2380 2390 2400  
GTTAAAATAC TATTGATGA CATTCAATTTC TCCTAGTTT TTTTTAGATA

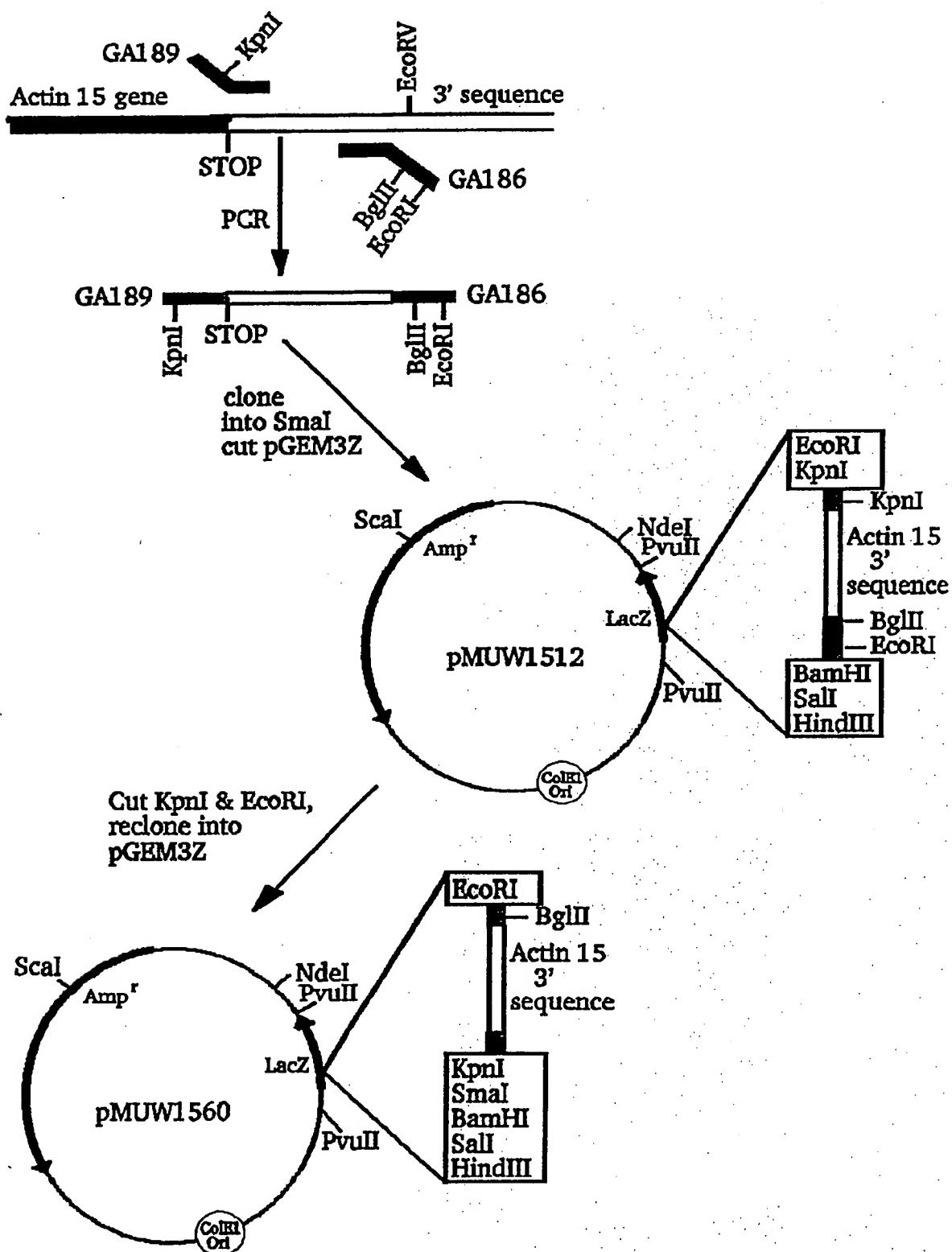
2410 2420 ClaI  
GATATAAAAA TAAATTGCCT AT

Fig. 11



24/30

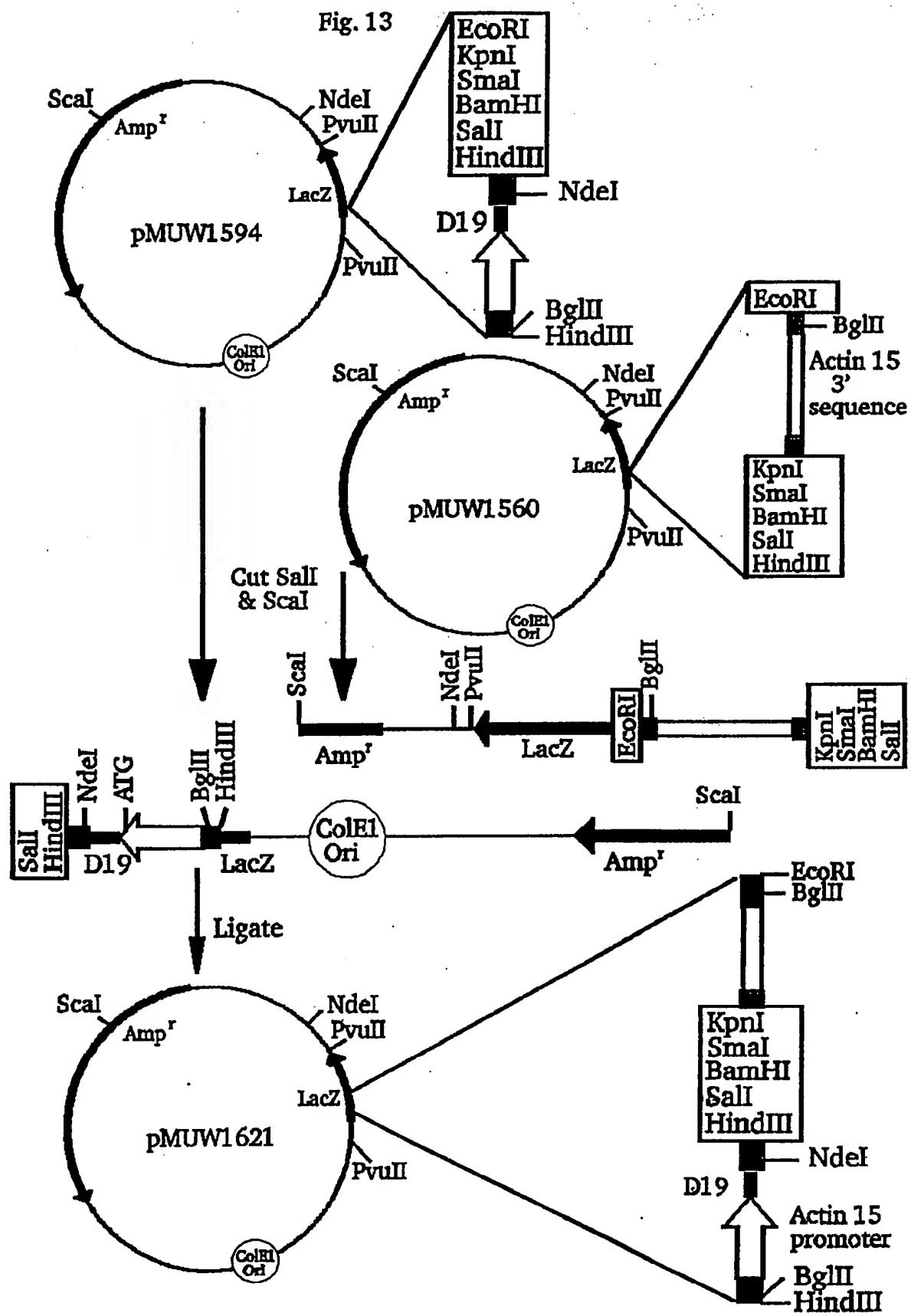
Fig. 12



SUBSTITUTE SHEET

25/30

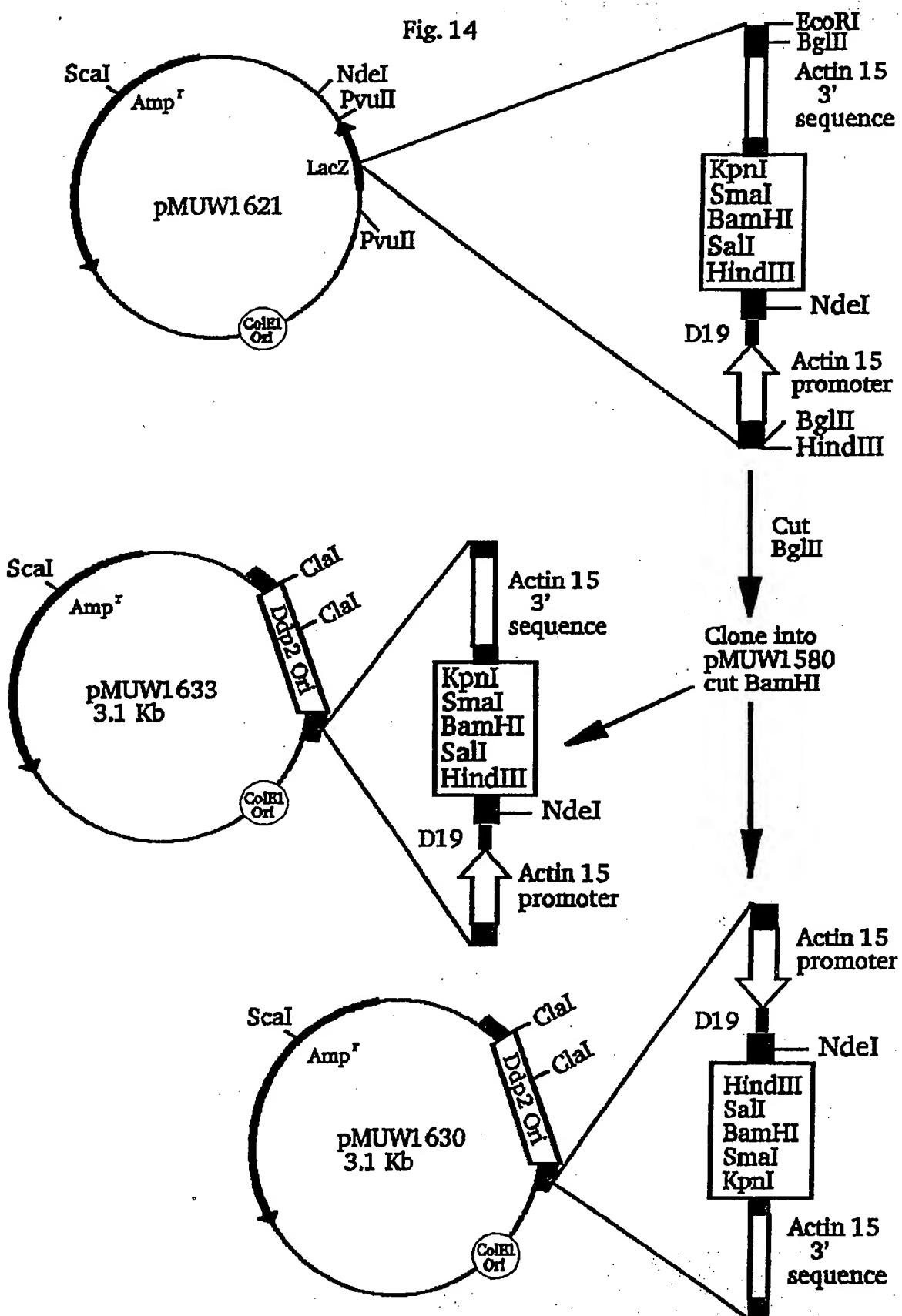
Fig. 13



SUBSTITUTE SHEET

26/30

Fig. 14



27/30

Fig. 15-1

ClaI 10 20 30 40 50  
 CGATAGGTGG CACTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTA  
 60 70 80 90 100  
 TTTTCTAAA TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCTG  
 110 120 130 140 150  
 ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT  
 160 170 180 190 200  
 TCCGTGTCGC CCTTATTCCC TTTTTGCGG CATTTCGCCT TCCTGTTTT  
 210 220 230 240 250  
 GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG  
 260 270 280 290 300  
 TGACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCATT  
 310 320 330 340 350  
 AGAGTTTCG CCCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT  
 360 370 380 390 400  
 CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGC AAGAGCAACT  
 410 420 430 440 ScaI 450  
 CGGTCGCCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG  
 460 470 480 490 500  
 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT  
 510 520 530 540 550  
 GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAAC  
 560 570 580 590 600  
 GATCGGAGGA CCGAAGGGAGC TAACCGCTTT TTTGCACAAAC ATGGGGGATC  
 610 620 630 640 650  
 ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA  
 660 670 680 690 700  
 AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGCCA CAACGTTGCG  
 710 720 730 740 750  
 CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA  
 760 770 780 790 800  
 TAGACTGGAT GGAGGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC  
 810 820 830 840 850  
 CTTCCGGCTG GCTGGTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG  
 860 870 880 890 900  
 GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA

28/30

Fig. 15-2

910	920	930	940	950
TCGTAGTTAT CTACACGACCG GGGAGTCAGG CAACTATGGA TGAACGAAAT				
960	970	980	990	1000
AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC				
1010	1020	1030	1040	1050
AGACCAAGTT TACTCATATA TACTTTAGAT TGATTAAAAA CTTCATTTTT				
1060	1070	1080	1090	1100
AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA				
1110	1120	1130	1140	1150
ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC CCGTAGAAAA				
1160	1170	1180	1190	1200
GATCAAAGGA TCTTCTTGAG ATCCTTTTT TCTGCGCGTA ATCTGCTGCT				
1210	1220	1230	1240	1250
TGCAAACAAA AAAACCACCG CTACCAGCGG TGGTTTGTGTT GCCGGATCAA				
1260	1270	1280	1290	1300
GAGCTACCAA CTCTTTTCC GAAGGTAACT GGCTTCAGCA GAGCGCAGAT				
1310	1320	1330	1340	1350
ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA				
1360	1370	1380	1390	1400
ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAACCT GTTACCAAGTG				
1410	1420	1430	1440	1450
GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTTGG ACTCAAGACG				
1460	1470	1480	1490	1500
ATAGTTACCG GATAAGGCCG AGCGGTCGGG CTGAACGGGG GGTTCGTGCA				
1510	1520	1530	1540	1550
CACAGCCCAG CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG				
1560	1570	1580	1590	1600
CGTGAGCTAT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA AGGC GGACAG				
1610	1620	1630	1640	1650
GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCACG AGGGAGCTTC				
1660	1670	1680	1690	1700
CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT TCGCCACCTC				
1710	1720	1730	1740	1750
TGACTTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATC				
1760	1770	1780	1790	1800
GAAAAACGCC AGCAACGCGG CCTTTTACG GTTCCTGGCC TTTTGCTGGC				

29/30

Fig. 15-3

1810	1820	1830	1840	1850	
CTTTGCTGG	CCTTGGATC	TACAAATTAA	TTAATCCCAT	CAAATCTTA	
1860	1870	1880	1890	1900	
AAAAAAAAAA	TGGTTTAAAA	AAACTTGGGT	TGGTTAATTA	TTATTTGAAA	
1910	1920	1930	1940	1950	
ATTTTAAAC	CCAAATTAAA	AAAAAAAAT	GGGATTCAAA	AATTTTTTT	
1960	1970	1980	1990	2000	
TTTTTTTTT	TTTTTTTTT	TTTTTTTTT	TTTTTTTCA	GATTGCATAA	
2010	2020	2030	2040	2050	
AAAGATTTT	TTTTTTTTT	TTTCTTATTT	CTTAAACAA	ATAAATTAAA	
2060	START	2080	2090	2100	
TTAAATAAAA	AATAAAAATG	AAATTCCAAC	ATACATTTAT	TGCATTATTA	
2110	2120	NsiI	HindIII	2140	2150
TCACTATTAA	CATACGCCAA	TGCATATGAA	AGCTTGCATG	CCTGCAGGTC	
2160	SmaI	KpnI	2180	2190	2200
GACTCTAGAG	GATCCCCGGG	TACCTAAATC	ATGAATGAAA	GTGCTTCACA	
2210	2220	2230	2240	2250	
TAAAAATAAT	AATAATAATA	TAACAATAAT	AATATTTAAA	TGTATAATAA	
2260	2270	2280	2290	2300	
AATTTAATTA	CTTTTTTTT	AATGGTTGTT	GATCTTATC	CGACCTTAAA	
2310	2320	2330	2340	2350	
AAAAAAAAAA	TAAAACCAAT	AGGCTATTGG	TTTTTTTTT	AATTGTTTTT	
2360	2370	2380	2390	2400	
TTATTTTTA	TTATTACTTT	AATTATCATT	TTTAAATTA	CAAAAAAAAT	
2410	2420	2430	2440	2450	
TAAAAATCCA	GATATTAAGG	TATTCGACT	AGTGCTTAA	CGTTAAAATT	
2460	2470	2480	2490	2500	
TGAAAAAA	AAAAAATTAA	TAATTTTACC	CTTTATGGGT	AAACGATTCT	
2510	2520	2530	2540	2550	
CACATATAAT	ACAATCTCCA	TGAAAAGATC	CGCTAGACGA	GCACAAATAT	
2560	2570	2580	2590	2600	
ATACTTTTTA	TTAAAAACGG	AGGTCAATTTC	AATACCTATT	GAAGAAATAA	
2610	2620	2630	2640	2650	
ATTTTTTTT	TTTTTTTTT	TGTCATGACA	CTTTTTTTT	TTTGTCAATGA	
2660	2670	2680	2690	2700	
CAGAATTGAA	AAAAACAGAA	AGTTATATAT	TTACCCCCCTT	TAATTTTTT	

30/30

Fig. 15-4

2710 2720 Clai 2740 2750  
TAÄAACTTTT GAAACTTTAG TAATAAGATC GATCTATACT TCAGTACGAA

2760 2770 2780 2790 2800  
CATAAATATG TATAAACCAA AAAAATTGAT TAAGATAAAG TTATATGTTT

2810 2820 2830 2840 2850  
GTATTTAATA AAATAGTTA GTTTAAAATT TTATATCATT TTTTAAAAAA

2860 2870 2880 2890 2900  
TGAAAATGTT TGAAAAAA AATTTTTTTT TTTTTTTCA ACGGGACGAT

2910 2920 2930 2940 2950  
GTAATATCAT ATATGATTCA AAATTAAAAG TTATTAACAA ATATGTAAAA

2960 2970 2980 2990 3000  
ATTATAAAAA ACTAACCTAG TTATAATTAC TTTCCCTCT TTTTTTTTT

3010 3020 3030 3040 3050  
TTTTTTGTC ATGACACTTT TTTTTTTTG TCATGACACT TTTTTTTAA

3060 3070 3080 3090 3100  
AAAAAAAAAA AAAAATGTTA AAATACTATT TGATGACATT CATTTCCT

3110 3120 3130  
AGTTTTTTT TAGATAGATA TAAAAATAAA TTGCCTAT

## INTERNATIONAL SEARCH REPORT

International Application No: PCT/AU 90/00530

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) 6

According to International Patent Classification (IPC) or to both National Classification and IPC

Int. Cl.<sup>5</sup> C12N 15/11, C12N 15/79, 15/80, C07K 13/00

## II. FIELDS SEARCHED

Minimum Documentation Searched 7

Classification System	Classification Symbols
IPC	DERWENT DATA BASES : WPAT, USPA, BIOTECHNOLOGY KEYWORDS : DICTYOSTELIUM

Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched 8AUST CLASS as above  
CHEM ABS using Keywords above

## III. DOCUMENTS CONSIDERED TO BE RELEVANT 9

Category*	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages 12	Relevant to Claim No 13
X	Plasmid Vol 20 (1988) (Barbara Leiting and Angelika Noegel) "Construction of an Extrachromosomally replicating transformation vector for <u>Dictyostelium discoideum</u> " pp 241-248. Whole document	(1,2,5,6,14) (17,22-24)
Y	Proc. Natl. Acad. Sci. USA, Vol 86 October 1989 Joseph L. Dynes and Richard A. Firtel (Molecular complementation of a genetic marker in <u>Dictyostelium</u> using a genomic library" pp 7966-7970, see page 7966 line 1-71 and page 7969 2nd paragraph - page 7970.	(1,2,5,6,14, 17,22-24)
A	Gene, Vol 39, (1985) Wolfgang Nellen & Richard A. Firtel "High copy number transformants & co transformation in <u>Dictyostelium</u> " pp 155-163.	

(continued)

* Special categories of cited documents: 10	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step
"E" earlier document but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

## IV. CERTIFICATION

Date of the Actual Completion of the International Search 8 February 1991 (08.02.91))	Date of Mailing of this International Search Report 20 February 1991
International Searching Authority Australian Patent Office	Signature of Authorized Officer <i>A. W. Bestow</i> A. W. BESTOW

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

A The Embo Journal Vol 2 No.4 (1983) Metz et al "Identification of an endogenous plasmid in Dictyostelium discoideum" pp 515-519.

## V. [ ] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE 1

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claim numbers ..., because they relate to subject matter not required to be searched by this Authority, namely:

2. [ ] Claim numbers , because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claim numbers .... because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4 (a):

## VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING 2

This International Searching Authority found multiple inventions in this international application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

[ ] The additional search fees were accompanied by applicant's protest.

[ ] No protest accompanied the payment of additional search fees.